Author:
Liu, Juntai

Title:
Engineering of polychromatic photosystems for expanded solar energy conversion

Engineering of polychromatic photosystems for expanded solar energy conversion

By Juntai Liu

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Life Science

World count:XXX

Abstract

Photosynthesis provides the mechanisms through which sunlight powers most of our biosphere. Either through direct application or inspiration, natural solar energy conversion strategies offer mankind potential solutions to impending energy and food crises through the exploitation of free solar power. However, the overall efficiency of photosynthesis is limited by a variety of factors including the selective light spectral coverage displayed by the choice of major pigments. Intriguingly, the complementary absorption profiles of chlorophyll-based photosystems and bacteriochlorophyll-based photosystems from oxygenic or anoxygenic phototrophs provide a pathway toward enhanced light capture across the photosynthetically-useful spectrum by synthetic biology. The research described in this thesis explores the effectiveness of a range of linking strategies to assemble bacteriochlorophyll-containing reaction centers (RC) and chlorophyll-containing light harvesting complexes (LHCs) into single polychromatic photosystems. Among the all biological strategies, a SpyTag/SpyCatcher linker offered the most effective way to form macromolecular “chimeras” between the RC from photosynthetic bacterium *Rhodobacter sphaeroides* and LHCs from *Arabidopsis thaliana*. Energy transfer from LHC to RC was confirmed both in solution and on an electrode and the current output of bacterial-RC photoelectrodes was shown to benefit from light capture by LHCs. In parallel, synthetic optically-active quantum dots (QDs) were shown to act as hubs for the self-assembly of LHC/RC/QD conjugates and to act as an energy bridge to augment direct LHC to RC energy transfer. A tight-binding interface between proteins and QDs was characterized. Based on the in-depth understanding of thermodynamics of this photosystem, it was found that energy flow within the tri-component conjugates could be tuned in a manner similar to natural photosystems and to a comparable level of efficiency. The project demonstrated the power of applying synthetic biology principles to the systematic redesign of natural photosynthesis and the expansion of solar energy conversion beyond the natural boundaries of living systems.

Acknowledgements

At first, I shall express my greatest gratefulness to my parents, my fiancée Jun Dong and other relatives for their unconditional and unreserved support.

Secondly, my sincere appreciation has to be given to staff and students at the University of Bristol across various institutes including Dr. Natalie Di Bartolo for providing XylE as a gift and all the people in the School of Biochemistry especially those I share office and laboratory space with on A floor in the Biomedical Sciences building. I also wish to thank Dr. Chris Wood and Prof. Dek Woolfson (School of Chemistry) for their assistance on the coiled coils work and the SpyCoil design; Dr. Fabio Parmeggiani (School of Biochemistry) for intellectual input on protein design; Dr. Benjamin Bode (School of Physics), Dr Majid Mosayebi and Prof. Tanniemola Liverpool (School of Mathematics) for their valuable suggestions on theoretical and computational work; Dr. Robin Corey and Dr. Deborah Shoemark for the advice on protein molecular dynamics; Mr. Jonathan Jenkins and Dr. Ahmad F.A. Sobri for sharing of *E. coli* strains and plasmids; Mrs. Judith Mantell and Dr. Chris Neal (Wolfson Bioimaging facility) for their warm help on TEM operation and data collection and the enthusiastic technician team on A-Floor in School of Biochemistry for their indispensable efforts in running the laboratory and making up hundreds of litres of growth medium. I also with to thank Dr. Liangfei Tian and Prof. Steven Mann (School of Chemistry); Prof. Ian Collinson (School of Biochemistry); Dr. Ross Anderson (School of Biochemistry); Dr. Paul Curnow (School of Biochemistry); Dr. Tom Oliver (School of Chemistry) and Dr. Adam Perriman (School of Cellular and Molecular Medicine) for sharing their equipment. I would like to thank Prof. Dek Woolfson and Prof. Nigel Savery for their suggestions as progression panel members.

I also want to thank our external collaborators and suppliers of materials including Prof. Mark Howarth (University of Oxford) for sharing the SpyTag/KTag/SpyLigase constructs; Prof. Roberta Croce and Mrs. Laura M. Roy (Vrije Universiteit Amsterdam) for sharing the LHCII construct to initialize the project; Dr. Vincent Friebe and Dr. Raôul Frese (Vrije Universiteit Amsterdam) for collaboration and their expertise on electrochemistry.

Moreover, I want to thank my supervisor Dr. Mike Jones for his dedication and supervision of this project.

At the end, without the funding from BBSRC&EPSRC Synthetic Biology Doctoral Training Centre (SynBio CDT) and a Faculty Fellowship University of Bristol, I will not have an opportunity to undertake the PhD. I would also like to give my acknowledgement to BrisSynBio, funding from which helped to sustain my research activity.

Author’s declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: .....  ..... DATE:.....27/06/2019.....

Table of Contents

1. Chapter 1 Introduction ---- 21
1.1. Powering life ---- 22
1.2. Global food and energy issues ---- 22
1.2.1. The food crisis ---- 22
1.2.2. The energy crisis ---- 24
1.2.3. World efforts for more and cleaner energy ---- 24
1.2.3.1. EU targets on CO_2 emission and renewable energy ---- 24
1.2.3.2. UK targets on CO_2 emission and renewable energy ---- 24
1.2.3.3. USA targets on renewable energy ---- 25
1.2.3.4. China's five-year plan on renewable energy ---- 25
1.2.4. Photosynthesis as a potential solution to global crises ---- 25
1.3. Photosynthesis ---- 26
1.3.1. Photosynthesis powers life ---- 26
1.4. Chlorophyll – the fundamental pigment for photosynthesis ---- 27
1.4.1. Chlorophyll *a*/*b* – oxygenic photosynthesis ---- 27
1.5. Chlorophyll *d*/*f* as an alternative ---- 28
1.5.1. Bacteriochlorophyll *a* – bacterial choice ---- 30
1.6. Reaction centers – the heart of photosynthesis ---- 30
1.6.1. Structure – sophisticated pigment-protein complexes ---- 30
1.6.1.1. PSII structure. ---- 30
1.6.1.2. PSI structure ---- 31
1.6.1.3. Bacterial RC structures ---- 31
1.6.2. Mechanism - converting solar energy into hot electrons and holes ---- 32
1.6.2.1. PSII – splitting water for electrons ---- 32
1.6.2.2. PSI – a dual carriageway for electron transfer ---- 34
1.6.2.3. The *Rba. sphaeroides* RC – a closed electron circuit ---- 34
1.7. Light harvesting systems ---- 37
1.7.1. Green plant light harvesting complex II ---- 37
1.7.2. Green plant light harvesting complex I ---- 38
1.7.3. Purple bacterial light harvesting complexes ---- 38
1.8. Organization of components in photosystems ---- 41
1.8.1. PSII/LHCII architecture ---- 41
1.8.2. PSI/LHCI supercomplexes and LHCII shuttling ---- 41
1.8.3. Rba. sphaeroides chromatophores ---- 43

1.8.4. Other LH structures - phycobilisomes and chlorosomes----43
1.8.5. Total energy conversion and efficiency----44
1.8.5.1. 100 TW of energy conversion (high) ----44
1.8.5.2. Biohybrid photoelectrochemical cells----44
1.8.5.3. Energy and quantum efficiencies ----45
1.8.5.4. Photosynthesis is not optimized to productivity ----45
1.8.5.5. Nature's limitations and strategies for improvement----45
1.9. Toward a bottom-up redesign of light capture in photosynthesis using a synthetic biology approach----46
1.9.1. QDs – artificial light harvesting units ----49
1.9.2. The QD/His-tag interface----49
1.9.3. Tuning QDs for photoprotein energy coupling----51
1.9.4. The arsenal of membrane protein compatible linkers ----51
1.9.4.1. SpyTag/SpyCatcher----52
1.9.4.2. *De novo* designed coiled coils----53
1.9.4.3. The split intein----54
1.9.4.4. SpyCoil – combining the merits of coiled coils and the Spy-system----55
1.10. Aims of this study----56
2. Chapter 2 Materials and Methods----58
2.1. Bacterial strains----59
2.1.1. Escherichia coli strains----59
2.1.2. Rhodobacter sphaeroides strains ----59
2.2. Bacterial growth, harvesting and storage ----60
2.2.1. Rhodobacter sphaeroides----60
2.2.2. Escherichia coli ----61
2.3. Expression vectors ----61
2.4. Molecular Biology----62
2.4.1. Preparation of plasmid DNA----62
2.4.2. Preparation of PCR products and digested linear DNA----62
2.4.3. Restriction digests and ligations----62
2.4.4. PCR amplification ----62
2.4.5. DNA transformation into chemically competent cells----62
2.4.6. DNA sequencing----63
2.4.7. Oligonucleotide and gene design and synthesis ----63
2.4.8. Site-directed mutagenesis ----64
2.4.9. Extensive modification of sequence by Gibson assembly ----64

2.4.10. Horizontal gene transfer from *E. coli* S17-1 to *Rhodobacter*. ---64
2.5. Protein expression in *Rhodobacter* and protein purification---65
2.5.1. General procedures ---65
2.5.2. Purification of His-tagged RCs---65
2.5.3. Purification of RCs lacking a His-tag ---66
2.5.4. Purification of His-tagged RC-LH1 complexes ---66
2.6. LHC expression in *E. coli* and *in vitro* reconstitution ---67
2.6.1. LHCII expression as inclusion bodies ---67
2.6.2. Pigment extraction ---68
2.6.3. LHCII reconstitution and purification---69
2.6.4. Spinach LHCII extraction and purification ---70
2.6.5. LHCI modification, refolding and purification. ---70
2.7. XylE expression and purification---70
2.8. eGFP/MBP soluble protein expression and purification ---71
2.9. Peptide synthesis and characterization ---71
2.9.1. Peptide synthesis ---71
2.9.2. Peptide purification and characterization---72
2.10. Polyacrylamide gel electrophoresis---72
2.10.1. SDS-PAGE ---72
2.10.2. Blue Native (BN) PAGE ---72
2.10.3. Western Blot ---73
2.11. Quantum yield determination ---73
2.11.1. QD and XylE/QD conjugate quantum yield ---73
2.11.2. LHCII quantum yield ---74
2.11.3. Quantum yields of other LHCII or LHCI---74
2.12. Transmission Electron Microscopy---74
2.12.1. Grid preparation ---74
2.12.2. Post imaging processing and analysis ---75
2.13. Dynamic light scattering---75
2.14. Sucrose density gradient ultracentrifugation---75
2.14.1. Sucrose cushion for membrane isolation ---75
2.14.2. Continuous sucrose density gradients for separation of RCs, LHCs and chimeras -75
2.14.3. Two-step sucrose density gradients for separation of proteins and conjugates ---76
2.15. Fluorescence titrations ---76
2.16. Charge separation in RCs---77

2.17. Molecular modelling ---- 78
2.17.1. Design of SpyCoil using ISAMBARD and Modeller ---- 78
2.17.2. Model construction with SCWRL and Modeller ---- 78
2.17.3. Molecular dynamics with GROMACS ---- 78
2.17.4. Analysis of surface coverage effects by Monte Carlo 2D packing of hard particles 79
2.18. Photochronoamperometry and EQE action spectra. ---- 79
3. Chapter 3 Mechanisms of Self-assembly and Energy Harvesting in Conjugates of Cadmium Telluride Quantum Dots and the *Rhodobacter sphaeroides* Reaction Centre ---- 80
3.1. Introduction ---- 82
3.2. Results ---- 84
3.2.1. Methodologies ---- 84
3.2.2. Mechanism of binding of RCs to QDs. ---- 84
3.2.3. Conjugation to QDs enhances RC charge separation. ---- 87
3.2.4. Energy is transferred from QDs to photochemically-inactive RCs. ---- 91
3.2.5. QD emission quenching is sensitive to spectral overlap. ---- 95
3.2.6. The morphology of RC/QD conjugates. ---- 97
3.2.7. Heterogeneity of actual RC:QD stoichiometry. ---- 99
3.2.8. Estimations of FRET efficiency and distance. ---- 106
3.3. Discussion ---- 110
3.4. Conclusions ---- 115
3.5. Supplementary Figures ---- 117
4. Chapter 4 Minding the Gap between Plant and Bacterial Photosynthesis for Solar Energy Conversion ---- 121
4.1. Introduction ---- 123
4.2. Results ---- 127
4.2.1. Assembly and characterisation of di-component LHCII/QD conjugates ---- 127
4.2.2. Assembly of $LHCII_H$/QD conjugates and evidence for $LHCII_H$ to QD energy transfer 133
4.2.3. Energy transfer efficiency and FRET distance for $LHCII_H$/QD conjugates. ---- 133
4.2.4. Architecture of $LHCII_H$/QD conjugates ---- 136
4.2.5. Architectures of tri-component conjugates ---- 137
4.2.6. Chl to BChl energy transfer in tri-component conjugates. ---- 139
4.2.7. Thermodynamic analysis of the composition of tri-component conjugates ---- 143
4.2.8. Population distribution and energy transfer ---- 150
4.2.9. Estimation of the efficiencies of direct and indirect energy transfer. ---- 150
4.2.9.1. Estimation of E_{Direct}. ---- 150
4.2.9.2. Estimation of $E_{Indirect}$ – effect of fraction of FRET-inactive conjugates. ---- 152

4.2.9.3. Estimation of $E_{Indirect}$ – determination of $E'_{LHCIItoQD}$. ---- 154
4.2.9.4. Estimation of $E_{Indirect}$ – determination of E'_{QDtoRC}.---- 156
4.2.10. Relative contributions of direct and indirect $LHCII_H \rightarrow RC_H$ energy transfer in tri-component conjugates ---- 159
4.3. Discussion ---- 161
4.4. Conclusions ---- 163
4.5. Supplementary Information ---- 164
4.5.1. Multiple independent binding model for $LHCII_H$/QD conjugates.---- 164
4.6. Supplementary figures ---- 165
5. Chapter 5 Polychromatic Solar Energy Conversion in Pigment-Protein Chimeras that Unite the Two Kingdoms of (Bacterio)Chlorophyll Photosynthesis ---- 177
5.1. Introduction ---- 179
5.2. Results ---- 182
5.2.1. Methodologies ---- 182
5.2.2. Solar energy conversion by unadapted photosystem components ---- 182
5.2.3. Design and production of components for chimeric photosystems ---- 187
5.2.4. Self-assembly of two-component RC-LHCII chimeras ---- 192
5.2.5. Self-assembly of two- and three-component RC-LHCI chimeras---- 195
5.2.6. Chl to BChl energy transfer in chimeras in dilute solution ---- 198
5.2.7. Functionality of chimeras on electrodes---- 201
5.2.8. Energy transfer efficiency in chimeras in solution---- 204
5.2.9. Energy transfer efficiencies on an electrode ---- 207
5.2.10. Effect of packing density on energy transfer efficiency for mixtures and chimeras---- 207
5.3. Discussion ---- 213
5.4. Conclusions ---- 215
5.5. Supplementary figures ---- 217
6. Chapter 6 Evaluation of Alternative Interface Components for Photosystem Construction. ---- 222
6.1. Introduction ---- 224
6.2. Results ---- 226
6.2.1. Evaluation of the use of coiled coils to form arrays of RCs ---- 226
6.2.1.1. Engineering and purification of double coiled-coil RCs (AR1 - AR4)---- 226
6.2.1.2. AR mutants show a low level of oligomerization at high salt concentration--- 228
6.2.1.3. Inducement of aggregation of the AR mutants by lowering ionic strength ---- 228
6.2.1.4. Array mutant RCs - summary ---- 231

6.2.1.5. Evaluation of heterodimeric coiled coils for formation of RC/LHCII heterodimers---------- 231
6.2.2. Design of heterodimeric coiled coil RC and LHCII constructs ---------- 231
6.2.2.1. Expression and purification of the CCA/CCB RCs and LHCIIs---------- 232
6.2.2.2. BN-PAGE screening of coiled-coil RC and LHCII pairs---------- 234
6.2.2.3. Characterization of the stoichiometry of 4K/E RC and LHCII oligomers--------- 236
6.2.2.4. Characterization the stoichiometry of 3.5 K-CC RC and 3.5E-CC LHCII aggregates ----------240
6.2.2.5. Conclusions regarding the use of coiled coils to assemble photoprotein oligomers ---------- 242
6.2.3. Evaluation of heterodimerization of RC/LHCII by the Cfa intein ---------- 242
6.2.3.1. Modification of RC and LHCII sequences with the Cfa split intein domains ---- 243
6.2.3.2. Heterodimerization of RC and LHCII with a split intein ---------- 243
6.2.3.3. Conclusions regarding the use of split inteins to assemble photoprotein oligomers ---------- 244
6.2.4. Design of SpyCoil and its utilization for mega-protein assembly---------- 246
6.2.4.1. Rationale behind the design of SpyCoil---------- 246
6.2.4.2. *In silico* design of SpyCoil ---------- 246
6.2.4.3. Molecular dynamics study of SpyCoil ---------- 247
6.2.4.4. Evaluation of the functionality of the designed SpyCoil linker ---------- 250
6.3. Discussion---------- 253
6.4. Conclusions---------- 255
6.5. Supplementary figures---------- 256
7. Summary and Outlook ---------- 266

List of Tables

Table 2.1 *E. coli* strains used in this work. ---------- 59

Table 2.2 *Rhodobacter sphaeroides* strains used in this work. ---------- 60

Table 2.3 Expression plasmids. ---------- 61

Table 3.1. Effect of linker length on the maximum number of RCs that can pack around a QD and overall conjugate diameter. ---------- 115

Table 5.1. Nomenclature and descriptions of pigment-protein complexes. --------- 192

Table 5.2. Apparent energy transfer efficiencies and associated parameters. ------ 208

Table 5.3. Apparent energy transfer efficiency on an electrode. ---------- 211

Table S5.1. Fits of P870 photobleaching and photon absorbance by RCs and LHCs. ---------- 224

Table S6.1. Highlight of the modification to the WT protein sequence. ---------- 268

List of Figures

Figure 1.1 Light powers life. ---------- 23

Figure 1.2 Pigments for light energy conversion. ---------- 29

Figure 1.3 RCs and their cofactor architectures. ---------- 33

Figure 1.4 Examples of RCs used for biohybrid devices with electron flow schemes. --- 36

Figure 1.5 Plant and purple bacterial antenna complexes. ---------- 39

Figure 1.6 RC-LHC supercomplexes and examples of natural biological control mechanisms. ---------- 42

Figure 1.7 Bottom-up construction of a hybrid Chl/BChl photosystem. ---------- 48

Figure 1.8 QDs filling the gap. ---------- 50

Figure 1.9 Peptidyl interfaces for chimera self-assembly. ---------- 53

Figure 3.1 Structure of the RC and optical properties of components. ---------- 85

Figure 3.2 Binding of RCs to QDs and quenching of QD emission. ---------- 87

Figure 3.3 Quenching of QD emission by His-tagged WT RCs. ---------- 88

Figure 3.4 Enhanced RC photobleaching in QD conjugates with and without carotenoid. ---------- 91

Figure 3.5 RC photobleaching at 870 nm in QD conjugates with and without carotenoid. ---------- 92

Figure 3.6. Quenching of QD emission by RCs lacking the P BChls. ---------- 94

Figure 3.7. Linearity of conjugate emission with concentration. ---------- 96

Figure 3.8. Quenching of QD emission by cofactor-exchanged RCs. ---------- 98

Figure 3.9. Dimensions of RC/QD conjugates. ---------- 100

Figure 3.10. Assessment of the fraction of RCs not bound to QDs. ---------- 102

Figure 3.11. Assessment of distribution in RC:QD stoichiometry. ---------- 105

Figure 3.12. Modelling of the distribution in RC:QD stoichiometry. ---------- 106

Figure 3.13. Distributions in RC:QD stoichiometry at different overall RC:QD ratios. ---- 107

Figure 3.14. Estimations of FRET efficiency and distance. ---------- 109

Figure 3.15. Determination of quantum yield of QDs. ---------- 110

Figure 3.16. Illustration of RC packing around a central QD. ---------- 114

Figure S3.1. Kinetics of dark recovery of strong P photobleaching at 870 nm. ---------- 119

Figure S3.2. Binding to QDs does not affect the kinetics of RC $P^{+}Q_{B}^{-}$ charge recombination. ---------- 120

Fig. S3.3. Emission spectra for VL157R RC:QD conjugates. ---------- 121

Figure S3.4. TEM images. ---------- 122

Figure 4.1. Pigment-protein structure and the mechanism of solar energy conversion. - 126

Figure 4.2. Tri-component conjugate design. ---------- 128

Figure 4.3. Binding of $LHCII_H$ to QDs. ---------- 130

Figure 4.4. Effect of His-tag position/length on binding of LHCII to QDs. ---------- 131

Figure 4.5. Quality control of refolded LHCII. ---------- 133

Figure 4.6. Determination of the quantum yield of $LHCII_H$. ---------- 134

Figure 4.7. Binding and energy transfer in $LHCII_H$/QD conjugates. ---------- 136

Figure 4.8. Architectures of protein/QD conjugates examined by TEM. ---------- 140

Figure 4.9. Energy transfer and charge separation in $LHCII_H/RC_H$/QD conjugates. ------ 142

Figure 4.10. $LHCII_H/RC_H$/QD conjugate controls. ---------- 144

Figure 4.11. Variance of conjugate composition with initial protein mixture. ---------- 146

Figure 4.12. Modelling of $LHCII_H/RC_H$/QD conjugate assembly. ---------- 147

Figure 4.13. Parameterization of $LHCII_H/RC_H$/QD conjugation modelling. ---------- 150

Figure 4.14. Extents of $LHCII_H/RC_H$/QD heterogeneity and FRET-inactive sub-populations. ---------- 151

Figure 4.15. Highlight of fraction of conjugates incapable of indirect energy transfer. --- 155

Figure 4.16. Energy transfer efficiencies at all tested $LHCII_H$:RC_H:QD ratios. ---------- 157

Figure 4.17. Correlations between E'_{QDtoRC} and $E_{DA,RC}$ and between E'_{QDtoRC} and $E'_{QDtoRC.}$ ---------- 160

Figure 4.18. Efficiencies and routes of energy transfer in $LHCII_H/RC_H$/QD conjugates. - 162

Figure S4.1. Deconvolution of emission spectra from $LHCII_H$/QD conjugates. ---------- 167

Figure S4.2. Determination of spectral overlap. ---------- 168

Figure S4.3. TEM of conjugates. ---------- 169

Figure S4.4. RC P870 photobleaching, dark recovery and fitted kinetics. ---------- 170

Figure S4.5. Kinetics of RC P870 photobleaching and dark recovery. ---------- 171

Figure S4.6. Energy transfer in tri-component conjugates. ---------- 172

Figure S4.7. Computed population heterogeneities for $LHCII_H/RC_H$/QD conjugates. ---- 173

Figure S4.8. Estimations of the sub-population of $LHCII_H/RC_H$/QD conjugates inactive in FRET. ---------- 174

Figure S4.9. Simulation of heterogeneity of $LHCII_H$/QD conjugate composition. ---------- 176

Figure S4.10. Final fraction distribution of $LHCII_H$/QD conjugate composition in simulations. ---------- 177

Figure S4.11. Sucrose gradient ultracentrifugation of $LHCII_H$/QD conjugates. ---------- 178

Figure 5.1. Component structures. ---------- 183

Figure 5.2. Mixing unadapted LHCII and RCs in solution. ---------- 186

Figure 5.3. Mixing unadapted LHCI and RCs in solution. ---------- 187

Figure 5.4. Photocurrents from electrodes coated with unadapted proteins. ---------- 188

Figure 5.5. Design of components for RC-LHC chimeras. ---------- 191

Figure 5.6. LHC adaptation and LHCII refolding. ---------- 193

Figure 5.7. Assembly of RC-LHCII chimeras. ---------- 195

Figure 5.8. Formation of covalent RC-LHCII chimeras. ---------- 196

Figure 5.9. Engineering of RC-LHCI chimeras. ---------- 198

Figure 5.10. Formation of covalent RC-LHCI chimeras. ---------- 199

Figure 5.11. Energy transfer in RC-LHCII chimeras in solution. ---------- 202

Figure 5.12. Energy transfer in RC-LHCI chimeras in solution. ---------- 203

Figure 5.13. Energy transfer in chimeras on an electrode. ---------- 205

Figure 5.14. Solar energy conversion in chimeras. ---------- 206

Figure 5.15. Schematics of adsorption of proteins on a roughened silver electrode. ---- 212

Figure 5.16. Packing of equimolar unadapted LHCII and RC at eleven percentage surface occupations. ---------- 214

Figure 5.17. Packing of LHCII#RC chimeras at eleven percentage surface occupations. ---------- 215

Figure 5.18. Estimated energy transfer efficiencies for WT mixtures and RC#LHCII chimera ---------- 217

Figure S5.1. Lamp spectra and EQE spectra recorded with a xenon lamp. ---------- 220

Figure S5.2. Purification of covalent RC#LHCs chimeras. ---------- 221

Figure S5.3. Emission quenching in RC#LHCII chimeras. ---------- 222

Figure S5.4. Emission quenching in RC#LHCI chimeras. ---------- 223

Figure 6.1. Constructs and quality of array RC mutants. ---------- 230

Figure 6.2. Sucrose gradients for assessing aggregate states of double coiled-coil RC mutants. ---------- 232

Figure 6.3. Response of aggregation state of mutant AR4 mutants to pH. ---------- 233

Figure 6.4. Response of AR3 RC to buffer ionic strength. ---------- 233

Figure 6.5. Constructs and quality of heterodimeric coiled coil RC mutants. ---------- 236

Figure 6.6 BN-PAGE of eight combinations of E/K coiled-coil RCs and LHCIIs to screen for heterodimers. ---------- 238

Figure 6.7. Sucrose density gradient analysis of 4K/4E coiled-coil RC/LHCII pairs. ----- 240

Figure 6.8. Analytical gel-filtration analysis of 4E/4K RC/LHCII pairs. ---------- 242

Figure 6.9. Effects of buffer/detergent conditions on heterodimer formation. ---------- 244

Figure 6.10. Cfa mutant constructs. ---------- 246

Figure 6.11. Construction of a RC/LHCII heterodimer using intein Cfa. ---------- 248

Figure 6.12. *In silico* design of SpyCoil.- ---------- 251

Figure 6.13. Molecular dynamic analysis of SpyCoil constructs. ---------- 252

Figure 6.14. Construct of designed SpyCoil proteins. ---------- 254

Figure 6.15. SDS-PAGE/WB analysis on SpyCoil eGFP dimerization with SpyCoil ----- 255

Figure S6.1. Original image of SDS-PAGE and Western blot. ---------- 259

Figure S6.2. Original image of SDS-PAGE and Western blot. ---------- 260

Figure S6.3. Original image of SDS-PAGE and Western blot. ---------- 261

Figure S6.4. Normalized fluorescence spectra of all coiled coil LHCII mutants. ---------- 262

Figure S6.5. Absorbance profile of sucrose density gradient from Fig. 6.9A.---------- 262

Figure S6.6. Absorbance profile of sucrose density gradients shown in Fig. 6.9B. ------- 263

Figure S6.7. Use of Cfa intein – original gel used in Fig 6.11B. ---------- 263

Figure S6.8. Use of Cfa intein – original gel used in Fig 6.11C. ---------- 264

Figure S6.9. Use of Cfa intein – original gel used in Fig 6.11D. ---------- 265

Figure S6.10. Use of SpyCoil – original gel used in Fig 6.15A. ---------- 266

Figure S6.11. Use of SpyCoil – original gel used in Fig 6.15BC. ---------- 267

List of abbreviations

A.	*Arabidopsis*
Amp	ampicillin
ATP	adenosine triphosphate
BChl	bacteriochlorophyll
BPhe	bacteriopheophytin
BUDE	Bristol University Docking Engine
CdTe	cadmium telluride
Chl	chlorophyll
CLAHE	contrast-limited adaptive histogram equalization
CS	charge separation
CV	column volumes
Cyt	cytochrome
DCM	4-dicyanomethylene-2-methyl-6-p-dimethylaminostyryl-4H-pyran
DDM	*n*-dodecyl β-D-maltoside
DIC	*N,N'*-diisopropylcarbodiimide
DLS	dynamic light scattering
DMF	dimethylformamide
DNA	deoxyribonucleic acid
DOPE	discrete optimized protein energy
DTT	dithiothreitol
E.	*Escherichia*
EDTA	ethylenediaminetetraacetic acid
EQE	external quantum efficiency
ET	energy transfer
Fd	ferredoxin
FMO	Fenna-Matthews-Olson
FNR	ferredoxin-$NADP^+$ reductase
FRET	Förster resonance energy transfer
FWHM	full width at half maximum
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HOBt	4-formyl-3-methoxyphenoxy)butanoic acid, 1-hydroxybenzotriazole
HPLC	high pressure liquid chromatography
HRP	horseradish peroxidase
IB	inclusion body
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan	kanamycin

LB	Luria-Bertani
LDAO	lauryldimethylamine N-oxide
LDS	lithium dodecyl sulphate
LDS-751	laser dye styryl–751
LED	light emitting diode
LHCI	light harvesting complex I
LHCII	light harvesting complex II
LHCs	light harvesting complexes
MALDI-TOF	matrix-assisted laser desorption ionization time-of-flight
MBP	maltose binding protein
MD	molecular dynamics
MPA	3-mercaptopropionic acid
$NAD(P)^+$	oxidized nicotinamide adenine dinucleotide (phosphate)
NAD(P)H	reduced nicotinamide adenine dinucleotide (phosphate)
Neo	neomycin
NPQ	non-photochemical quenching
nt	nucleotide
NTA	nitrilotriacetic acid
ODE	ordinary differential equation
OEC	oxygen evolving complex
OG	octyl β-D-glucopyranoside
PBS	phosphate buffered saline
PCR	polymerase chain reaction
Phe	pheophytins
PQ	plastoquinone
PSI	Photosystem I
psi	pounds per square inch
PSII	Photosystem II
Q_0	ubiquinone-0
QD	quantum dots
Rba.	*Rhodobacter*
RC	reaction centre
RMSD	root-mean-square deviation
RMSF	root mean square fluctuation
RNA	ribonucleic acid
rpm	revolutions per minute
SDS-PAGE	sodium dodecyl sulphate – polyacrylamide gel electrophoresis

Strp	streptomycin
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
TCEP	tris(2-carboxyethyl)phosphine
Tch.	*Thermochromatium*
TEM	transmission electron microscopy
Tet	tetracycline
Tet^R	tetracycline resistance
TMAO	trimethylamine N-oxide
Tris	tris(hydroxymethyl)aminomethane
TW	terrawatts
UQ	ubiquinone
WB	Western blot
WT	wild-type
YFP	yellow fluorescent protein

Chapter 1

Introduction - Synthetic Biology meets Photosynthesis

1.1. Powering life

Life on Earth can be divided into autotrophs and heterotrophs depending on the approach to obtaining energy. Autotrophs such as plants, algae and many bacteria either convert sunlight directly through photosynthesis (**Fig. 1.1** left) or derive energy from inorganic molecules in their environment through chemosynthesis. Heterotrophs such as humans, on the other hand, have to consume organic carbon created by other organisms for their livelihood (**Fig. 1.1** right). Some organisms such as the purple bacterium *Rhodobacter (Rba.) sphaeroides* can self-sustain through photosynthesis or grow on external nutrients and therefore can act as both an autotroph and heterotroph (**Fig. 1.1** middle). Alongside a limited contribution from reduced inorganic molecules from the Earth's interior, for example sources of reduced sulphur that are used by some chemoautotrophs such as archaea,[1] solar radiation is the main source of energy available to the Earth, powering a most extraordinary out-of-equilibrium chemical system – life.

1.2. Global food and energy issues

1.2.1. The food crisis

As we approach the end of the second decade of this new century, mankind is facing increasing challenges relating to energy and food production caused by growing populations and increasing living standards. By 2050, food demand is predicted to increase by 60%-120% compared to current provision[2]. However, the unavoidable shrinking of land available for farming as the result of growing populations and climate change caused by carbon release is leading to an urgent situation that will require fundamental changes to agricultural and energy production. Accordingly, research groups around the world are working towards improving photosynthesis yields by targeting individual components through genetic modification[3,4]. As an example, progress has been seen in research on various algae that have a better energy conversion efficiency than conventional plants and have the potential to be alternative food sources[5].

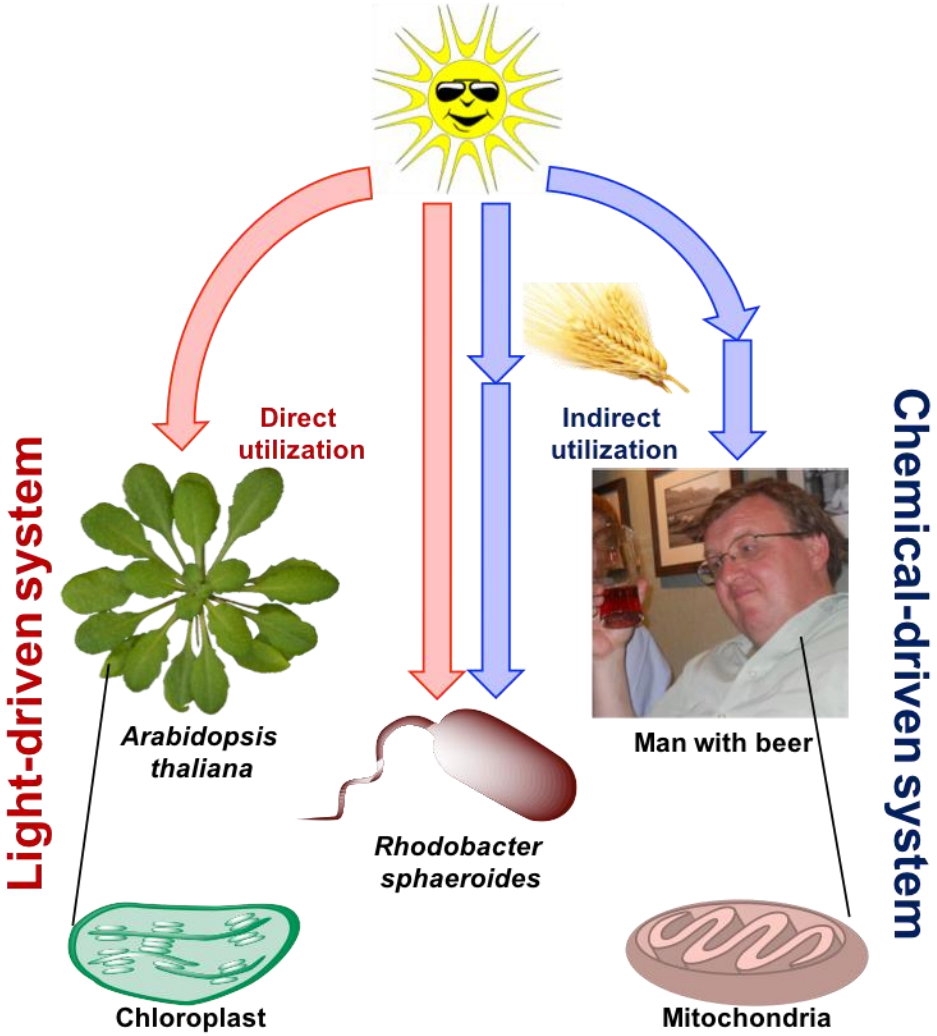


Figure 1.1 Light powers life. The majority of life forms on earth can either directly utilize sunlight (red arrows) or obtain energy indirectly via other chemical sources generated from solar energy (blue arrows). The plant *Arabidopsis thaliana* is an example of living system directly energized by sunlight through photosynthesis that takes place in chloroplasts. Humans (e.g. man with beer) are an example of indirect utilization of sunlight through chemical energy conversion within mitochondria. Some organisms such as *Rhodobacter sphaeroides* can grow in either way.

1.2.2. The energy crisis

Human civilisation consumed energy at a rate of around 18 TW(Terra Watts) in 2015,[6] most of which (>85%) came from coal, oil and natural gas[7]. This heavy dependence on energy dense fossil fuels has led to an increase of 30% in the amount of CO_2 in atmosphere in 2018 compared to the 1950 level and the current concentration of CO_2 has broken a new record since 650,000 years ago[8]. The so-called global warming from greenhouse gases has imposed a deterioration in the quality of our daily life represented by increasing extreme temperatures, droughts, floods, intense hurricanes and super typhoons (e.g. Mangkhut, which affected Hong Kong and China in 2018)[9–11]. Against this background of concern over fossil fuel dependency and climate change, it has been predicted that mankind's energy demand will increase to 27-30 TW by 2050 and to 43-46 TW by 2100[12,13].

1.2.3. World efforts for more and cleaner energy

1.2.3.1. EU targets on CO_2 emission and renewable energy

Regarding the energy issue, it was scheduled in the European Commission Energy Roadmap 2050 that 75% of total energy and 97% of electricity should be generated through renewable sources by the middle of the 21st century[14]. This report also highlighted goals for tackling climate change by decarbonisation that expects carbon release to be less than 5%-20% of the 1990s level[14]. Achieving this with no expansion of nuclear power generation means an increased dependence on renewable energy such as wind and biomass/solar energy. In addition, a revolution in renewable techniques will be necessary to improve energy conversion efficiencies and storage capacities.

1.2.3.2. UK targets on CO_2 emission and renewable energy

An increasing dependence on renewable energy was also reported in the recent 2018 UK Energy In Brief,[15] with low-carbon sources contributing 18% of total UK energy supply compared to 9.4% in 2000. With supply from nuclear power being static over this period the growth has come largely from electricity generation from wind, solar and biomass energy, which in 2017 was 15% greater than that from nuclear power, in contrast to being only 13% of that from nuclear power in 2000[15]. The UK and European Union have committed to a reduction in carbon emissions of >80% by 2050[16]. Both the UK and EU have also emphasized the importance of a revolution in renewable energy generation to reduce the demand on crops for bio-energy production, so that land can be freed for food production[14,17]. It was proposed that photovoltaic solar cells need to be improved to harvest more energy and connected to the

current framework or used in local areas to provide sufficient energy supply[14]. Under the circumstance of no substantial expansion of nuclear energy, solar energy, among all the renewable energy options, is the most promising primary source for both fuel and electricity generation and comprises the largest part of renewable energy sources.

1.2.3.3. USA targets on renewable energy

The USA has set an ambitious goal for its future energy industry of 80% coming from renewable sources[18]. A National Renewable Energy Laboratory report also stated in an energy roadmap for the USA that 50% of renewable electricity will come from wind or solar photovoltaics by 2030[18]. Some states such as California have achieved 32% of their electricity generation from renewable sources, of which solar energy occupies 36%, offering the largest contribution to renewable energy production. The installation of photovoltaics grew exponentially nation-wide in the USA compared with other sources of renewable energy such as wind, geothermal, hydro and biomass. Again, solar radiation was considered to be the most promising source of renewable energy.

1.2.3.4. China's five-year plan on renewable energy

China has the biggest solar power infrastructure, generating at the rate of 0.043 TW in 2015 with its carbon emission profile changing mostly because of production structure changes and efficiency gains[19]. In the 13th Five Year Plan on energy development, the role of solar energy was emphasized to fulfill the decarbonization target with an increasing dependence on non-fossil energy from 12% in 2015 to 15% by the end of 2020. It was stated that the facility capacity of solar power needed to be increased by 21.2% annually, much higher than other renewable sources such as hydro (2.8%) and wind (9.9%).[20]

1.2.4. Photosynthesis as a potential solution to global crises

There is great interest in how the study of natural photosynthesis can address potential future crises such as energy and food supply. Approaches that are being pursued include improving the efficiency or resilience of natural photosynthesis *in vivo*, applying features of natural photosynthetic structures and mechanisms to the design of synthetic materials for solar energy conversion, and using photosynthetic pigments and proteins as components in biohybrid devices for solar energy conversion and related technologies. All approaches depend on an understanding of how nature powers life using sunlight.

1.3. Photosynthesis

1.3.1. Photosynthesis powers life

Photosynthesis is a natural procedure to covert solar energy to chemical energy[21]. It stores the captured photon energy in the form of stable carbohydrates and so produces biomass[22]. Green plant and algal photosynthesis are the dominant forms, involving the reductive fixation of carbon dioxide into more complicated carbohydrate products with water as the electron donor and solar radiation as the energy supplier. Green plant photosynthesis is generally divided into two parts, the light reactions and the dark reactions. During the light reactions, plants absorb photon energy and lock-in the energy in the form of NADPH (reduced nicotinamide adenine dinucleotide phosphate) and ATP (adenosine triphosphate). Then, in the dark stage, ATP and NADPH are used to fix and reduce CO_2 to stable carbohydrates. Plant photosynthesis is one of the vital contributors to the maintenance of our current biosphere and to sustain the food chain and also our survival. Other photoautotrophic organisms, especially some bacteria, use electron donors other than water or cycle electrons within the photosystem, and also contribute significantly to global elemental cycles (e.g. carbon, oxygen, nitrogen)[23].

The form of photosynthesis found in plants, algae and cyanobacteria is termed oxygenic photosynthesis, as oxygen is produced from the oxidation of water during the light reactions. The photosynthetic apparatus is based around two membrane-embedded photosystems, named Photosystem I (PSI) and Photosystem II (PSII), that work in series to pass electrons from water to $NADP^+$. A variety of bacteria perform an alternative anoxygenic photosynthesis involving a single type of photosystem that has structural and compositional similarities to either PSII or PSI[24,25]. These anoxygenic phototrophs include purple bacteria (phylum Proteobacteria), green non-sulphur bacteria (phylum Chloroflexi) and a species from the phylum Gemmatimonadetes (*Gemmatimonas* sp. AP64[25]), which have PSII-like (or type-II) photosystems, and green sulphur bacteria (phylum Chlorobi), heliobacteria (phylum Firmicutes) and a species from the phylum Acidobacteria (*Candidatus* Chloracidobacterium thermophilum[24]) which have PSI-like (or type-I) photosystems. Links of functionality and structure between RCs from very different species suggest a common ancestor (see for recent review[26,27]).

Irrespective of the type of photosynthesis, the photosystems that facilitate solar energy conversion are large integral membrane proteins in which light harvesting complexes (LHCs) surround a central reaction centre complex (RC). The role of the light-harvesting or “antenna” pigment-protein complexes is to absorb solar energy and pass the energy of the resulting

excited electronic state to the RC. The cofactors of the RC then facilitate a photochemical separation of electrical charge across the photosynthetic membrane, initiating a cascade of electron transfer reactions among additional components of a linear or cyclic electron transfer chain that are coupled to proton translocation across the photosynthetic membrane.

The emergence of photosynthesis is undoubtedly the most important evolutionary event in the history of life[28]. The great oxidation of the Earth's atmosphere was the consequence of ancient cyanobacterial oxygenic photosynthesis that reduced atmospheric CO_2 from 3.6% to 0.02% and increased the oxygen level from nearly zero to about 21%[22,29]. This fundamentally changed the Earth's environment from reductive to oxidative and these cyanobacteria are still one of the most important contributors to the global carbon cycle, accounting for 20% to 30% of global photosynthetic productivity[30].

1.4. Chlorophyll – the fundamental pigment for photosynthesis

1.4.1. Chlorophyll *a*/*b* – oxygenic photosynthesis

Chlorophyll (Chl) *a* was once believed to be the only chlorophyll involved in oxygenic photosynthesis and is the most abundant chlorophyll-type pigment on Earth[31]. Many photosynthetic organisms, from cyanobacteria and algae to higher plants have adopted a Chl *a*-only photosystem and use it to harvest and convert solar energy to sugar and split water into oxygen, electrons and protons, whilst Chl *b* is mainly used for light harvesting in photosynthetic eukaryotes. The structure of Chl *a* is shown in **Fig. 1.2A** with the differences to other (bacterio)chlorophylls highlighted in colour. Tuning by the protein environment enables green phototrophs to use Chl *a* not only for light harvesting along with Chl *b*, but also for the implementation of photochemical charge separation in the very heart of photosynthesis.

The redox properties of Chl *a* are also key to its use in photosynthesis. In PSII the P680 Chl *a* can support an oxidizing potential that is sufficiently high to remove an electron from a tyrosine side chain and hence water. In PSI the initial Chl *a* electron acceptor has a sufficiently low reducing potential to drive a cascade of electron transfer reactions that end with the reduction of $NADP^+$. Hence Chl *a* only can operate at both the extreme oxidizing and extreme reducing limits of photosynthesis.

1.4.2. Chlorophyll *d*/*f* as an alternative

Recently, it was found that Chl *d* and Chl *f* are used for far-red light oxygenic photosynthesis in place of the dominant Chl *a* that defines the “red-limit” of green photosynthesis[32]. Chl *d* has the methyl group at the C3 position substituted by a more hydrophilic formyl group (**Fig. 1.2A**), while Chl *f* possesses a formyl group at its C2 position instead of the vinyl group in Chl *a*. These pigments can absorb light at longer wavelengths than the 700 nm limit set by Chl *a*, with Chl *d* extending absorption to 740 nm and Chl *f* providing absorption even above 760 nm[33]. Based on the discovery of these pigments, they are believed to be the result of adaptation of oxygenic phototrophs to an environment with limited visible light such as in a deep water niche[34].

Chl *d* and Chl *f* aid in the use of far-red light by either replacing Chl *a* in photoproteins to alter the light harvesting pattern or acting as the primary electron donor with a reduced redox potential generated through charge separation in both Photosystem II and Photosystem I. In the most studied Chl *d*-containing cyanobacterium, *Acaryochloris marina*, 97% of the chlorophyll was Chl *d* with only a few Chl *a* in the RC mixed with Chl *d* [35].

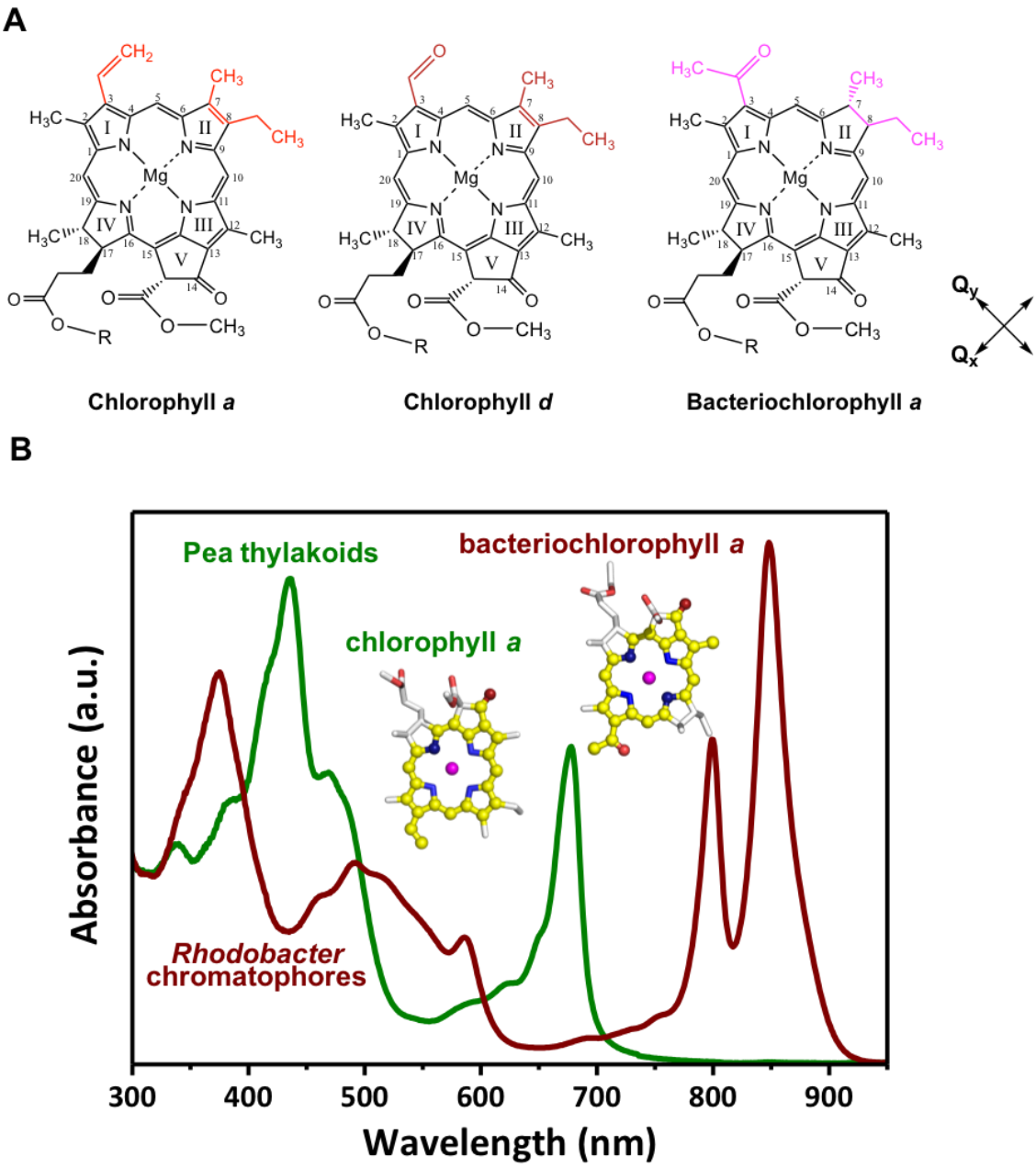


Figure 1.2 Pigments for light energy conversion. (A) Differences that affect π-conjugation are highlighted in colour: red – chlorophyll *a*; orange – chlorophyll *d*; purple – bacteriochlorophyll *a* (Figure prepared with the help from Dr. Vincent Friebe from Vrije Universiteit Amsterdam[31]). The directions of the Q_y and Q_x transition dipoles are indicated by arrows that cross at the central magnesium. Most differences are at ring carbons 3, 7 and 8 as labelled. R represents the hydrocarbon side chain. (B) Absorbance spectra of pea thylakoid membranes and the photosynthetic membrane of *Rba. sphaeroides*, demonstrating their complementary solar energy harvesting. The former one contained mostly chlorophylls and the latter one contained mostly bacteriochlorophylls with carotenoid molecules providing additional coverage of spectrum.

1.4.3. Bacteriochlorophyll *a* – bacterial choice

Bacteriochlorophyll *a* (BChl *a*) is the main photosynthetic pigment in the majority of anoxygenic photosynthetic bacteria including green sulphur bacteria, purple bacteria and filamentous anoxygenic phototrophs[36]. It is the key pigment that conducts the photochemical reactions and light harvesting function, and so resembles the function of Chl *a* in oxygenic phototrophic organisms[37]. The macrocycle of BChl *a* is different from Chl *a* at (1) the acetyl group at C-3 position on ring I and (2) reduction at ring II (double bond to single bond between C7 and C8) (**Fig. 1.2A**). By these differences, BChl *a* has a smaller and more asymmetric conjugated system, and hence a modified transition energy landscape relative to Chl *a*. This endows BChl *a* with the ability to absorb light at near-infrared or near-ultraviolet wavelengths that are outside of the Chl *a* absorption range from 400 nm to 700 nm. Therefore, a BChl *a*-based photosystem is optically not in direct competition with photosystems from oxygenic phototrophs. This complementarity is illustrated in **Fig. 1.2B** which compares the absorbance spectra of pea thylakoids and *Rba. sphaeroides* photosynthetic membranes. The ability to utilize light above 700 nm is particularly interesting because a large amount of solar energy arrives on Earth above this wavelength limit that green plants, algae and cyanobacteria cannot absorb. Pairs of BChl *a* also form the "special pair" where photochemical charge separation is initiated and interacting BChls also play a light harvesting role in bacterial photosynthesis[38]. Other types of BChl are also found among anoxygenic prokaryotes, including BChl *b* in many species of purple bacteria and BChl *g* which is the primary photosynthetic pigment in Heliobacteria[36].

1.5. Reaction centers – the heart of photosynthesis

All organisms that engage in chlorophyll-based photosynthesis contain one or two types of RC pigment-proteins that use sunlight to power redox reactions. This section looks at the structures and mechanisms of the three most heavily studied RCs.

1.5.1. Structure – sophisticated pigment-protein complexes

1.5.1.1. PSII structure.

The structure and mechanism of PSII have been extensively reviewed (see for recent examples [39–41]). The core pigments of PSII are Chl *a*, or Chl *d* in some cases[33] (**Figure 1.2A**). The structural model shown in Fig. **1.3A** is based on a high resolution X-ray crystal structure of PSII from the cyanobacterium *Thermosynechococcus vulcanus* (3WU2[42]), the central

pigment composition of which is conserved in the PSII of algae and green plants as well as other cyanobacteria. Besides four weakly-interacting Chl *a* that include the primary electron donor P680, two pheophytins (Phe_{D1}/Phe_{D2} – demetallated Chls) and two plastoquinones (PQ - Q_A/Q_B) are associated with the core domain formed from a heterodimer of D1 and D2 proteins. These cofactors are arranged around an axis of C_2 symmetry in two trans-membrane chains or branches, a structural feature also seen in other RCs. A tyrosine (Tyr_Z) and a Mn_4O_5Ca oxygen evolving complex (OEC) are present close to P680 and are essential for the water splitting reaction. In the wider PSII complex, which includes the CP43 and CP47 light harvesting complexes, there are 31 Chl *a* in addition to the central pigments conducting the photochemistry that efficiently distribute energy absorbed directly or received from additional surrounding light harvesting complexes.

1.5.1.2. PSI structure

The structure and mechanism of PSI have also been extensively reviewed (see for recent examples [39,40,43]). The structural model of PSI shown in Fig. **1.3B** is based on a high resolution X-ray crystal structure of PSI from *Pisum sativum* (4XK8[44]). Charge separation chains are formed by six Chl *a*, with two forming a special pair P700 (**Fig. 1.3B**, left, red). There are also two quinones and a Fe_4-S_4 iron-sulphur centre (Fx) in a heterodimer formed by the main core subunits PsaA and PsaB. Of note, the B_2/A_2 Chl *a* are coordinated by PsaB and PsaA in a way that is opposite to the other cofactors on each electron transfer chain. Two additional FeS centers (F_A and F_B) are incorporated into the electron transfer chain within a small protein domain called PsaC. Similar to PSII, additional Chl *a* perform light harvesting and energy transfer functions. In the case of the structure used in **Fig. 1.3B** from cyanobacteria, there are about 90 peripheral Chl *a* of which 79 are bound within the PsaA/PsaB core, significantly higher than count of peripheral Chl *a* in PSII. A recent study has shown that Chl *f* with a red-shifted absorption can replace the function of Chl *a* in PSI as the core light harvesting antenna or conversion of energy in the core electron transfer chain[33].

1.5.1.3. Bacterial RC structures

X-ray crystal structures have been determined for type-II RCs from the purple bacteria *Blastochloris viridis*[45], *Rba. sphaeroides*[46,47] and *Thermochromatium* (*Tch.*) *tepidum*[48] and the type-I RC from *Heliobacterium modesticaldum*[49]. Here the focus will be on the structure of the most extensively studied RC, it is the complex from *Rba. sphaeroides*[38,50–53], which was used as an experimental material throughout the work described in this thesis. As can be seen from the structural model in **Fig. 1.3C**, which is based on X-ray crystal structure 1PCR[54], this RC is

formed from three polypeptides that are named H (or more formally PuhA), M (PufM), and L (PufL). PufL and PufM form a heterodimeric structure in the membrane that acts as a scaffold for the electron transfer cofactors, whilst PuhA insulates the quinone electron acceptors from the aqueous phase with LH1-koncked out mutant, while in native environment RC was insulated with LH1 ring. The RC cofactors consists of four BChl *a* (P_A/P_B, B_A/B_B), two bacteriopheophytin (BPhe) *a* (H_A/H_B), two ubiquinones (Q_A/Q_B), a carotenoid and an iron atom (**Fig. 1.3C**). No extra BChls or other pigments are present in this type of bacterial RC to provide additional absorption, and as the *Rba. sphaeroides* RC is much smaller and has fewer pigmented components than PSII or PSI it is possible to assign absorption bands to every pigment and their energy state as shown in **Fig. 1.3D**, endowing this bacterial RC as a good target for protein engineering.

1.5.2. Mechanism - converting solar energy into hot electrons and holes

1.5.2.1. PSII – splitting water for electrons

As outlined above, PSII is capable of oxidizing water to oxygen in the OEC with the energy input from sunlight, with a quantum yield (electrons removed per photon absorbed) of nearly unity.[39,41,43] Photochemical charge separation occurs only along the left cofactor branch shown in **Fig. 1.3A** and **1.4A**) which leads to the Q_A PQ. The initial charge separation is complex, with the possibility of more than one initial charge separated state among the four Chls and Phe_{D1}, but mechanisms converging on the electron being localized on Phe_{D1} and the highly-oxidizing cation ($P680^+$) being located on the P_{D1} Chl *a*. The excited electron is then funneled to the PQ at the Q_B site through Q_A with a progressively reduced energy potential. The electron hole localized on P_{D1} receives an electron from a water bound in the Mn_4O_5Ca OEC via Tyr_Z (**Fig. 1.3A** and **1.4A**). The final energy stored in PQ is about 60% of the energy incident on P680, higher than either PSI or the bacterial RC (**Fig. 1.4A** and **1.4C**). Electron holes are stably stored at the OEC which can exist in five redox states (S_0 – S_4) and the life-time of the

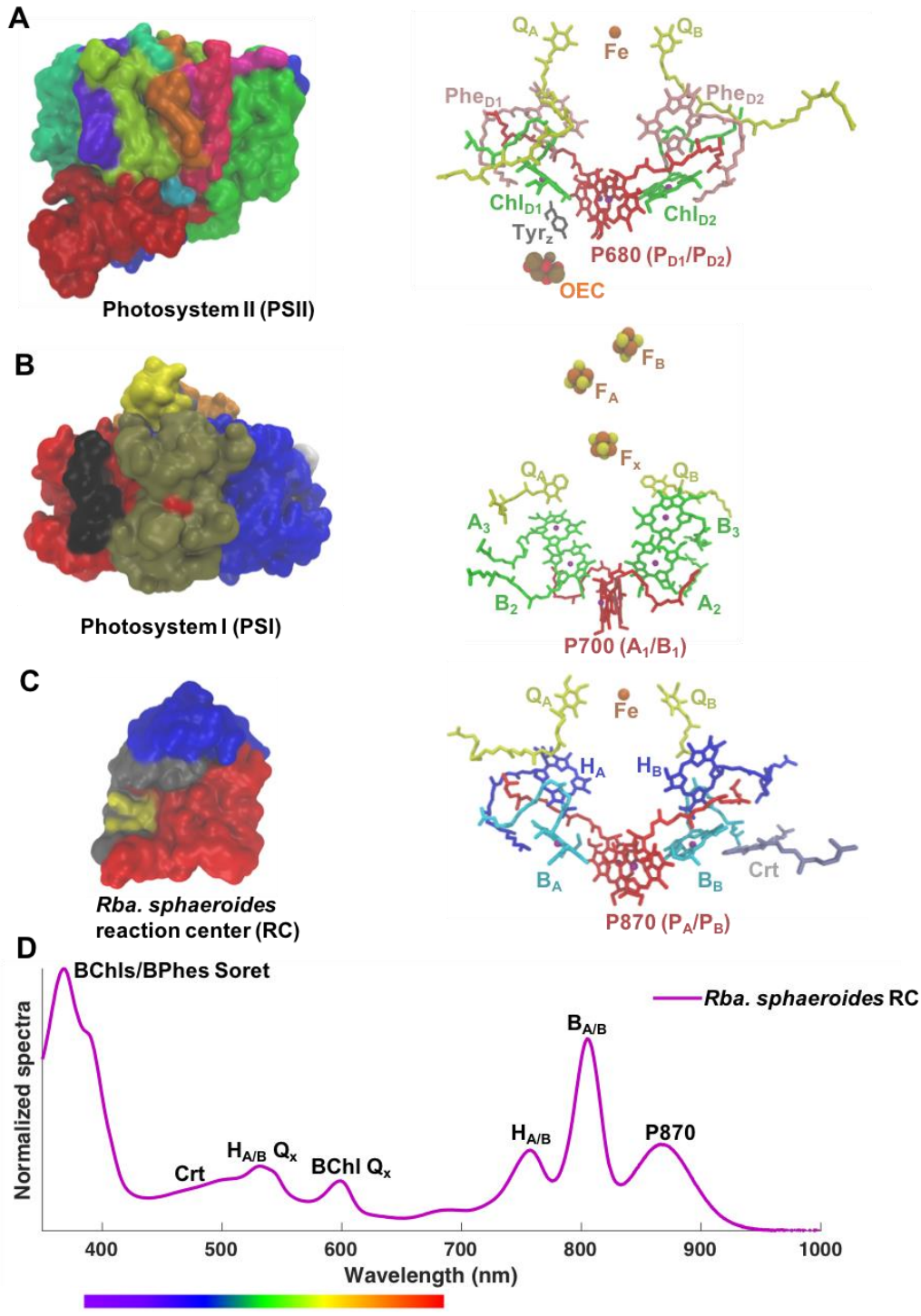


Figure 1.3 RCs and their cofactor architectures. (A) Proteins of a PSII monomer and cofactors of the PSII RC. Key: OEC – Mn_4O_5Ca cluster; Tyr_z – redox-active tyrosine; P680 and $Chl_{D1,D2}$ - Chl *a*; $Phe_{D1,D2}$ – Phe *a*; $Q_{A/B}$ – plastoquinone. (B) Proteins of a PSI monomer and cofactors of the PSI RC. $A_{1,2,3}/B_{1,2,3}$ – Chl *a*. $Q_{A/B}$ – phylloquinone; $F_{X,A,B}$ – Fe_4S_4 cluster. (C) Proteins of the *Rba. sphaeroides* RC and its cofactors. P870 and $B_{A,B}$ – BChl *a*; $H_{A,B}$ – BPhe *a*; $Q_{A/B}$ – ubiquinone; Crt – carotenoid. (D) Absorbance spectrum of the *Rba. sphaeroides* RC and the assignment of bands to pigments.

trans-PSII radical pair can be from 30 s to hours depending on the redox states of the OEC and Q_B site[55]. The energy received by PQ will be transferred via a membrane-embedded cytochrome (cyt) b_6f complex to an extra-membrane electron carrier such as plastocyanin, which is used as the electron donor to PSI.

1.5.2.2. PSI – a dual carriageway for electron transfer

In PSI it is thought that the initial charge separation takes place on either cofactor branch between two monomeric Chls (B_2/A_3 or A_2/B_3 in **Fig. 1.3B**) with a second step of hole transfer to the P700 Chl pair[39,40,43]. The electron lost from the resulting $P700^+$ is replenished by plastocyanin reduced by the b_6f complex (**Fig. 1.4A**). Different from PSII and the purple bacterial RC, electrons can pass along both branches because each quinone does not undergo double reduction/protonation but rather passes a single electron to the F_X FeS centre which sits on the C_2 symmetry axis (**Fig. 1.3B** right). Each generated electron will finally pass through two more FeS centres and on to an attached ferredoxin and the ferrodoxin-$NADP^+$ reductase (FNR), producing NADPH as reducing power to fix CO_2 to reduced sugar molecules (**Fig. 1.4A** right). Overall, PSI uses sunlight to reduce $NADP^+$ to NADPH and plastocyanin as an electron donor, and the internal energy conversion efficiency is about 0.52 (**Fig. 1.4C**). The charge pair $P700^+F_B^-$ has a half-life of around 0.3 s[56].

1.5.2.3. The *Rba. sphaeroides* RC – petite power pack

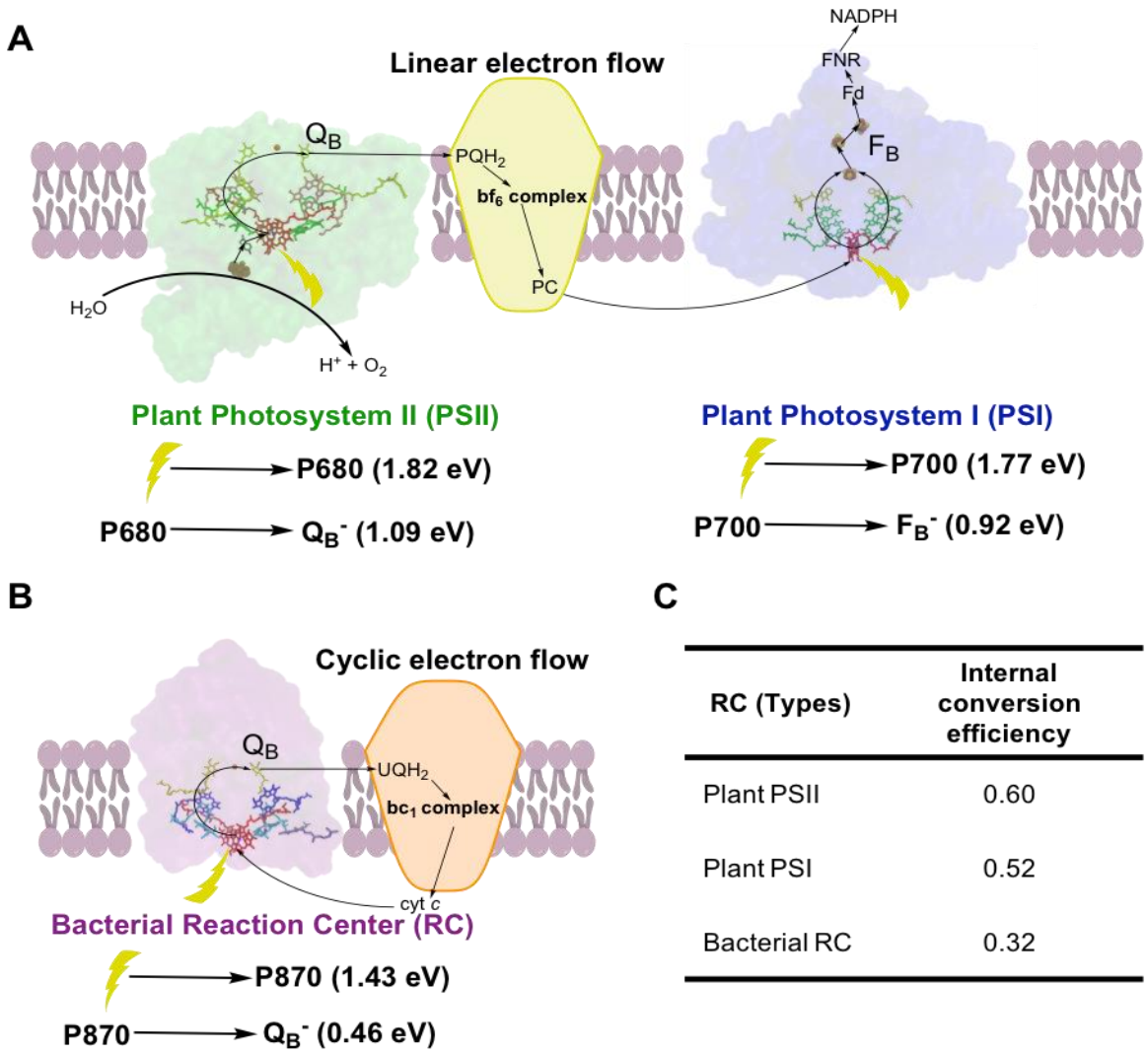
Instead of having a terminal electron donor and acceptor as in the plant/algal photosystem, the *Rba sphaeroides* photosystem does not require sacrificial redox-active reagents to complete its reaction cycle. A circuit of light-driven electron transfer is accomplished between the RC and a cyt bc_1 complex, with ubiquinone and a cyt c_2 as mediators from the RC to the bc_1 complex and back, respectively (**Fig. 1.4B**). The RC does not translocate protons but the cyt bc_1 complex operates a Q-cycle, such that the light-powered flow of electrons is coupled to proton translocation. The resulting proton gradient is used to generate ATP and assisted generation of NAD(P)H through a process called reverse electron flow in which reduction of the intra-membrane ubiquinone pool results in NAD^+ reduction by NADH dehydrogenase (Complex I).

Charge separation in the *Rba. sphaeroides* RC, one of the proteins used extensively in the work described in this thesis, has been studied very extensively and is well understood[38,51–53,57]. This in-depth characterisation, employing a very wide range of biophysical techniques including ultrafast spectroscopy, has been assisted by the determination of X-ray crystal

structures for the RC, beginning over 30 years ago, the ability to separate the RC from the surrounding light harvesting system, and the relatively uncongested absorbance spectrum of the isolated RC (**Fig. 1.3D**). Regarding the latter, the most insight has been obtained from the so-called Q_y region of the absorbance spectrum between 700 nm and 950 nm where the RC has three absorbance bands. The band at ~870 nm is attributable to the special pair of BChl *a* (called P870 or just P) which are excitonically-coupled, the band at ~802 nm to overlapping transitions from the two monomeric BChl *a* (B_A and B_B) and the band at ~760 nm to overlapping transitions from the two BPhe *a* (H_A and H_B). The four BChls also contribute to a Q_x absorbance band at ~600 nm and the two BPhes to a Q_x absorbance band at ~540 nm. All six bacteriochlorins contribute to the Soret absorbance band between 300 and 420 nm. The single RC carotenoid (when spheroidenone) produces a broad absorbance band between 450 nm and 600 nm.

Charge separation begins from the first singlet excited state of the P870 special pair (P870*), and proceeds in the sequence $P870^+B_A^-$, $P870^+H_A^-$, $P870^+Q_A^-$, and $P870^+Q_B^-$ with lifetimes of around 3-5 ps, 1 ps, 200 ps and 1 µs for the four steps[38,51–53,57]. The quantum efficiency of these initial charge separation events is close to 100% because at each stage the forward charge separation reaction is much faster than the competing reverse reaction or recombination to the ground state. $P870^+$ is re-reduced by a water-soluble cyt c_2 on the periplasmic side of the membrane in a few µs (**Fig. 1.4B**). There is then a second charge separation to doubly reduce the ubiquinone at the Q_B site, accompanied by a double protonation, and a second cyt c_2 is oxidized. This produces the substrates, QH_2 and two oxidized cyt c_2, for the cyt bc_1 complex. On the first turnover of the RC the final stored energy is 0.32 of the initial energy arriving on P870, with $P870^+Q_B^-$ state being stable for about one second in isolated RCs (**Fig. 1.4C**)[58,59].

In wild type (WT) *Rba. sphaeroides* an extensive light harvesting system delivers excited state energy to the RC and the Q_y absorbance band of P870 coincides with the strong absorbance band of the BChls of LH1 at ~875 nm. This means that energy enters the RC from LH1 mainly through energy transfer to P870. However, relevant to the work described in this thesis, in isolated RCs it is also possible to directly excite the monomeric BChls B_A and B_B and the BPhes H_A and H_B. If this occurs then energy is passed in a few hundred femtoseconds to P870 to start charge separation[60]. In addition, it is known that direct excitation of B_A can also produce charge separation by alternative mechanisms not starting from P870* (initial reactions $B_A^* \rightarrow B_A^+H_A^-$ or $B_A^* \rightarrow P870^+B_A^-$)[61].



C

RC (Types)	Internal conversion efficiency
Plant PSII	0.60
Plant PSI	0.52
Bacterial RC	0.32

Figure 1.4. Examples of RCs used for biohybrid devices with electron flow schemes. (A) Plant photosynthetic electron flow starting from PSII, where electrons are taken from water, yielding protons and O_2, and are then delivered to PQ (plastoquinone) at the Q_B site via one cofactor branch only. The reduced PQH_2 transfers the electron to plastocyanin via the cyt b_6f complex, which then donates the electron to PSI. Electron flow in PSI can take both branches, reaching the terminal Fe-S cluster (F_B), where it is passed to Fd (ferredoxin), FNR (ferredoxin-$NADP^+$ reductase) and finally $NADP^+$ which is reduced to NADPH. (B) The energized electron in the *Rba. sphaeroides* RC is delivered to a UQ (ubiquinone) at the Q_B site via the left branch only. Reduced UQH_2 passes electrons to cyt c_2 via the cyt bc_1 complex and the cyt c_2 recovers the charge at the RC P870. (C) Internal energy conversion efficiencies of the usual extractable electron pools by photovoltaic or photochemical devices versus starting energy of the primary electron donor. The internal conversion efficiency was calculated based on the division product of the final charge separated state to the initial excited primary electron donor. The difference of energy to ground states are shown in eV. In (A) and (B) black arrows show the

electron (energy) flow direction. Yellow lightning arrows represent absorbed photons that excite the RCs to trigger charge separation and sequential electron transfers.

1.6. Light harvesting systems

A wide variety of pigment-proteins contribute to solar energy harvesting in photosynthesis. This section concentrates on those most relevant to the work described in this thesis.

1.6.1. Green plant light harvesting complex II

Light harvesting complex II (LHCII) is the most abundant component in photosynthetic membranes from green plants[62–66]. The energy absorbed by LHCIIs can be transferred to PSII with a nearly 100% efficiency. LHCII usually exists as a trimeric combination of the proteins Lhcb1, Lhcb2 and Lhcb3 (**Fig. 1.5A**), with the pigment configuration of each monomeric domain being identical. Each monomeric unit of LHCII contains 8 Chl *a*, 6 Chl *b* and 3 carotenoids (**Fig. 1.5A**). A Chl *a* cluster with red-shifted absorbance (named Chl 610/611/612 in the X-ray crystal structure) is considered to be the emitting site following internal fast energy transfer from the other LHCII Chl and carotenoid pigments.

In addition to light harvesting, LHCII plays an important role in protection against excessive excitation in strong sunlight. Single molecule fluorescence measurements have unveiled blinking of LHCII emission intensity presumably due to protein dynamics, which is associated with the very important photoprotection function of LHCIIs[67]. There are several explanations for the energy dissipation pathway through which LHCII performs the so-called non-photochemical quenching (NPQ) such as internal structure changes especially focusing on zeaxanthin/violaxanthin cycle or inter-complex relaxation in aggregated LHCIIs dissociated from PSII[68–71].

In 1990, Paulsen and co-workers described the recombinant expression of an LHCII apoprotein in *Escherichia* (*E.*) *coli* and *in vitro* refolding of the holoprotein using purified Chl and carotenoid pigments[72]. This opens a route to genetic modification of LHCII, including the introduction of additional structural features that could interfere with LHCII function in the native organism. This approach was used in the work on adapted LHCII complexes described in this thesis.

1.6.2. Green plant light harvesting complex I

Light harvesting complex I (LHCI) is associated with PSI and is present as stable heterodimers of the proteins Lhca1/Lhca4 (**Fig. 1.5B**) and Lhca2/Lhca3[44,73–75]. The four LHCI subunits share a very similar protein scaffold to LHCIIs, which renders to them a somewhat similar pigment configuration. However the pigment composition of LHCI consists of dominantly Chl *a* and this produces a red-shifted emission above 720 nm. In LHCI heterodimers, Lhca4 and Lhca2 have been reported to be responsible for the red and far-red emission from a Chl *a* 603-609 pair (**Fig. 1.5B**). In contrast to LHCII hardly losing its "blue-peak" around 682 nm[67], LHCI constantly switches emission from an LHCII-like peak to a red form with the peak of emission shifted to longer wavelengths by more than 30 nm[76]. Their far-red fluorescence output could be a better candidate for energy transfer to bacterial RCs because of better dipole-dipole resonant interaction for energy exchange. LHCI proteins can also be expressed in *E. coli* and LHCI heterodimers can be refolded *in vitro*, facilitating protein engineering.

1.6.3. Purple bacterial light harvesting complexes

In all purple photosynthetic bacteria the RC is associated with a light harvesting complex denoted as LH1. In many species there is also a dominant peripheral LH2 antenna, and X-ray crystal structures have been determined for the LH2 proteins from *Rhodopseudomonas acidophila*[77] and *Rhodospirillum molischianum*[78] showing a protein architecture consisting of closed hollow rings (**Fig. 1.5C**). In *Rhodopseudomonas acidophila* LH2 consists of concentric rings of 9 copies each of 2 types of single membrane-span protein (α and β) that hold in place two BChls rings. One ring consists of 9 discrete BChls that have a characteristic absorption band at 800 nm, while the second ring contains 18 overlapping BChls that have a characteristic absorption band at 850 nm (**Fig. 1.5C**). The LH2 in *Rba. sphaeroides* has a similar structure.

The LH1 antenna forms a larger ring that surrounds the RC and the resulting RC-LH1 complex can be stably purified as a so-called "core complex". High-resolution structures have been determined for RC-LH1 complexes from *Tch. tepidum*[79] (**Fig. 1.5D**) and *Blastochloris viridis*[80], and a medium resolution structure has also been reported for the RC-LH1 from *Rba. sphaeroides*[81]. In the last case the RC-LH1 exists as a dimer that includes an intrinsic membrane polypeptide called PufX. To accommodate the RC, LH1 is larger than the LH2 hollow ring and in *Rba. sphaeroides* is formed from 14 copies of α/β pairs that coordinate 28

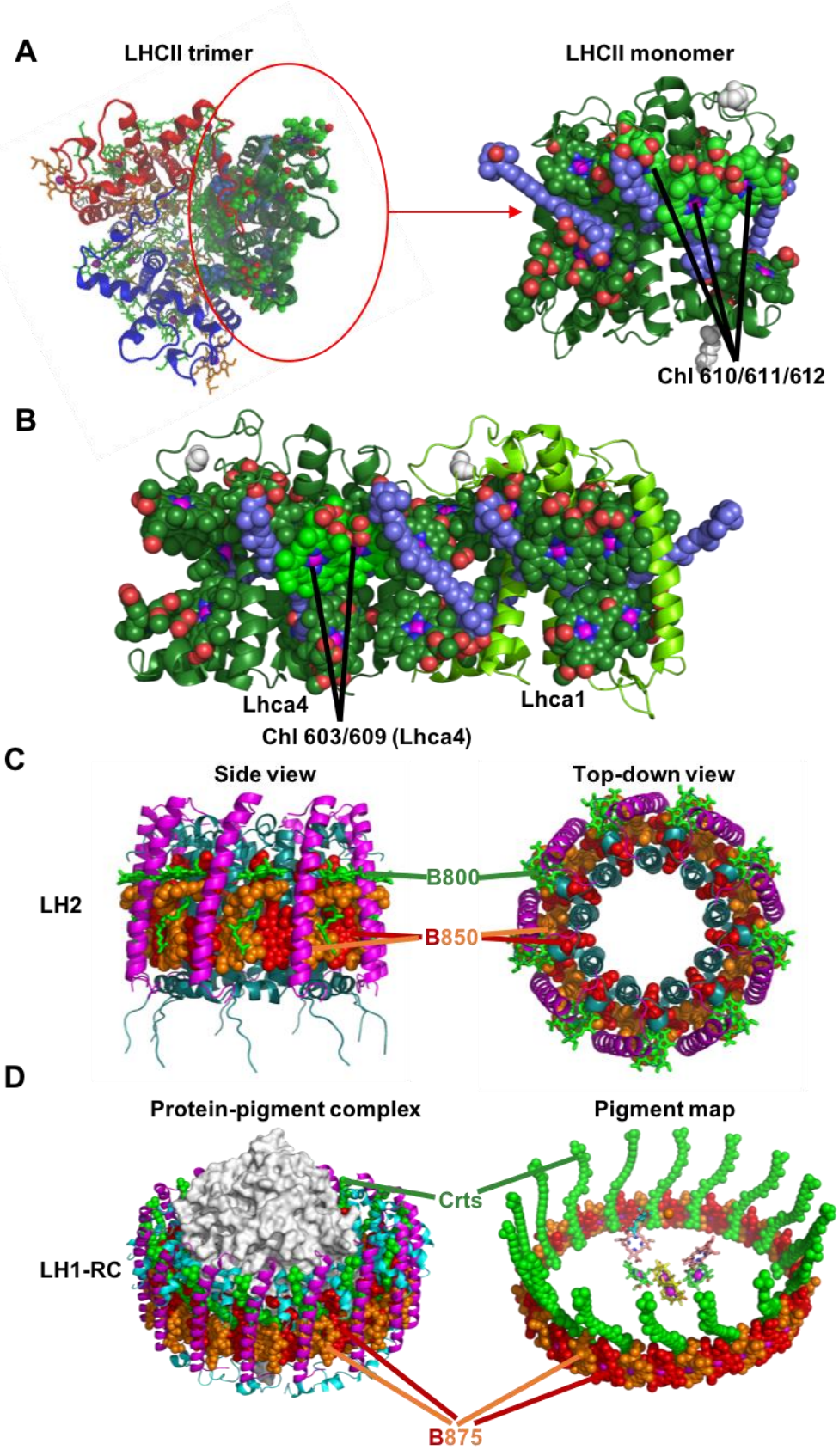


Figure 1.5 Plant and purple bacterial antenna complexes. (A) Plant LHCII mainly exists as as a trimer (right). Each monomer (left) binds 8 Chl *a* (green carbons), 6 Chl *b* (green carbons)

and 3 carotenoids (blue carbons). The red ellipse indicates an LHCII monomer and the low energy emitting Chl *a* cluster (610/611/612) is highlighted with light green carbons. The structure is PDB entry 2BHW for LHCII from *Pisum sativum*.[65] (B) Plant LHCI exists as heterodimer of Lhca1/4 (shown) or Lhca2/3 (not shown). Lhca1 accommodates 12 Chl *a* (green carbons), 2 Chl *b* (green carbons) and 4 carotenoids (blue carbons). Lhca4 consists of 11 Chl *a* (green carbons), 4 Chl *b* (green carbons) and 3 carotenoids (blue carbons). The red-most pigment cluster (Chl 603/609) on Lhca4 which the emitting terminus of the Lhca1/4 dimer is is highlighted with light-green carbons. The structure is PDB entry 4XK8 for PSI-LHCI from *Pisum sativum*.[75] (C) The *Rhodopseudomonas acidophila* LH2 antenna is a closed cylindrical structure formed from nine pairs of α and β proteins that scaffold 9 B800 BChls (green) and 18 B850 BChls (alternating red and orange). LH2 also contains carotenoids (not shown). The structure is PDB entry 1NKZ[82]. (D) In RC-LH1 complexes the LH1 antenna surrounds a central RC with up to 16 pairs of α and β proteins. Up to 32 LH1 B875 BChls (alternating red and orange) and 32 carotenoids (in *Rba sphaeroides*) surrounds the cofactors of the RC. The pigment map of LH1 shows that the B875 BChl ring is in the same plane as the RC P870 primary electron donor and also the B850 BChls of LH2, endowing a fast and efficient energy transfer from LH2 to LH1 and finally to the RC. The structure is PDB entry 3WMM for the RC-LH1 from *Tch. tepidum*[83] which has only one carotenoid per α/β pair.

BChls and 28 carotenoids.[84] The ring of overlapping BChl *a* gives an absorption band around 875 nm and acts as the donor in the flow of energy to the enwrapped RC (**Fig. 1.5D**). If PufX is removed then the ring of LH1 becomes closed through the addition of two extra α/β pairs, four BChls and four carotenoids, forming a monomeric RC-LH1 structure similar to that seen in *Thermochromatium tepidum* (**Fig. 1.5C** bottom panel)[85].

In *Rba. sphaeroides* the RC, LH1 and LH2 complexes can be expressed independently of one another, and removal of their structural genes from the genome can produce strains such as DD13 that lack any BChl complexes, and to which individual proteins can be added back[86]. The work described in this thesis made extensive use of so-called RC-only strains where genetically modified RCs are expressed in the absence of LH1 or LH2[87], greatly simplifying their purification and enabling extensive modification.

1.7. Organization of components in photosystems

1.7.1. PSII/LHCII architecture

The PSII core complex is formed from the D1/D2 heterodimeric RC and closely-associated CP43 and CP47 LHCs, and exists as a dimer in grana membranes (**Fig. 1.6A** left). This PSII core dimer is associated with four copies of trimeric LHCII and two copies of each of the minor LHCII monomeric proteins CP24, CP26 and CP29[88]. This supercomplex is called $C_2M_2S_2$ where C is the PSII core, M is a moderately strongly bound LHCII trimer and S is a strongly bound LHCII trimer (**Fig. 1.6A** left). Recently, a structure of the $C_2M_2S_2$ supercomplex was resolved by cryo-EM, which unveiled many effective natural strategies for photosynthesis[66]. For example, there are two potential energy pathways from the antenna complexes to the RC either via a S-LHCII red Chl *a* cluster or mediated by the minor complex CP26. On the other hand, M-LHCII can donate energy via either S-LHCII or a combination of CP29 and CP24. Beside the energetic interfacing, the CP24, CP26 and CP29 minor components also serve a vital role in orientating the major trimeric LHCII with PSII in the correct configuration to ensure efficient energy transfer. Knock out of these minor complexes abolished PSII-LHCII association and arrested growth of *Arabidopsis* (*A.*) *thaliana*[89], indicating the crucial role of the minor antenna complexes.

1.7.2. PSI/LHCI supercomplexes and LHCII shuttling

Plant PSI is multi-component complex and LHCIs are normally found to be attached to one side of each PSI, forming a belt (**Fig. 1.6A** right)[44]. These complexes are separated from PSII-enriched domains and are located at the membrane margin between grana and stroma membranes or at stroma lamellae. Despite X-ray crystal structures identifying plausible energy transfer pathways that include a few up-hill energy transfer steps, the trapping efficiency of LHCI to PSI is nearly unity (97.5%) as is the trapping efficiency from LHCII when it is attached to PSI (96%)[90], probably the highest efficiencies among the discussed antenna complexes with an associated RC. PSI has attracted wide interest for potential photovoltaic applications due to its better stability compared to PSII and its stronger light absorptivity compared to bacterial RCs[91].

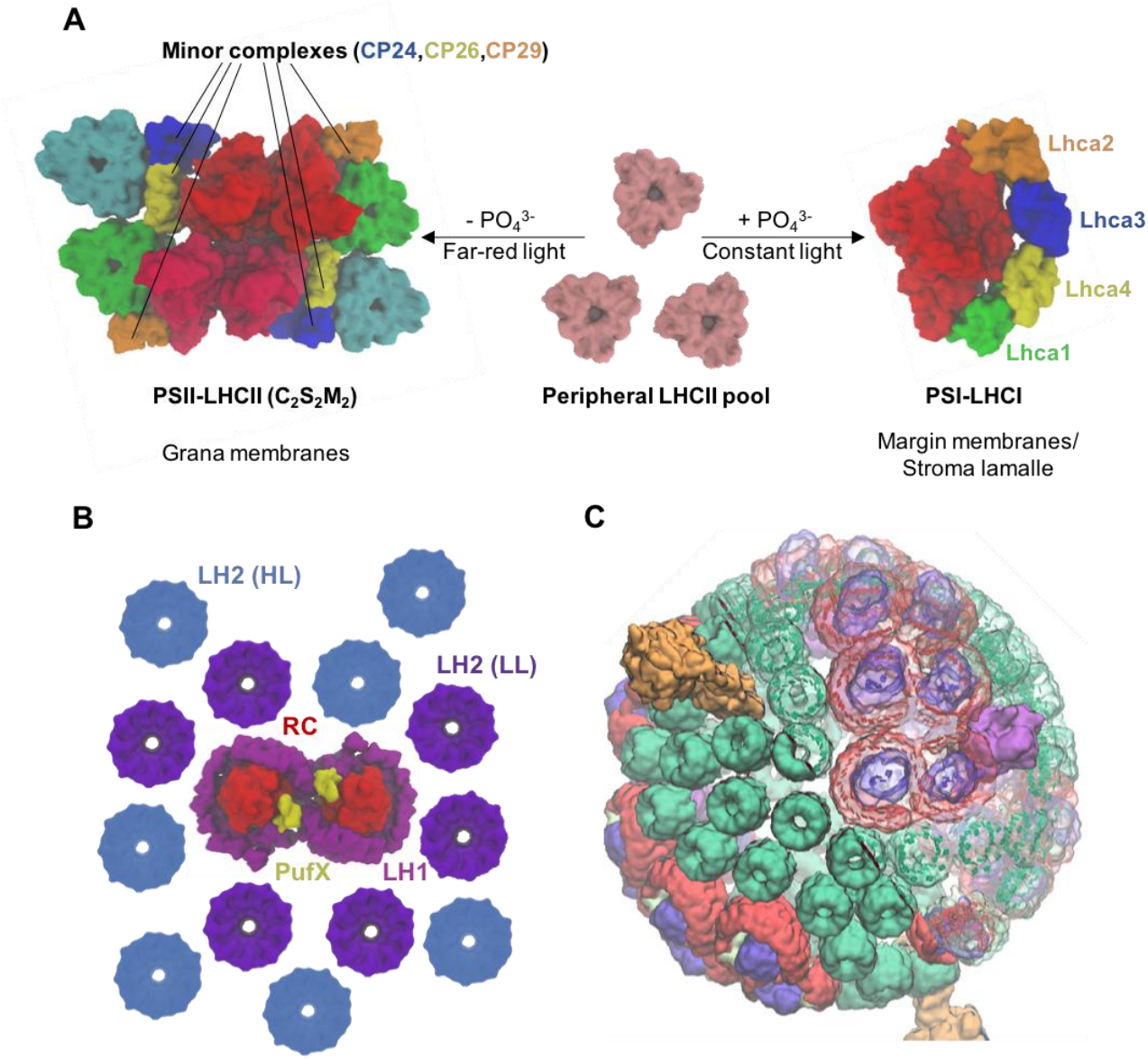


Figure 1.6 RC-LHC supercomplexes and examples of natural biological control mechanisms. (A) Architectures of PSII-LHCII and PSI-LHCI supercomplexes. LHCIIs bound to PSII strongly are shown in green and moderately are shown in cyan. The structure also contains with the assistance of minor LHCII complexes CP24 (blue), CP26 (yellow) and CP29 (orange). A PSI-LHCI monomer has single copies of the PSI RC and Lhca1/4 and Lhca2/3 heterodimeric complexes. Peripheral LHCIIs associates with the PSI RC from another side. Peripheral LHCIIs are allocated to core PSII-LHCII or PSI-LHCI supercomplexes depending on light-induced phosphorylation state changes. (B) Arrangement of bacterial RC-LH1-PufX core complexes and LH2 antenna complexes in the *Rba. sphaeroides* membrane. Under high light (HL) conditions more LH2 (blue) is synthesized compared to low light (LL) conditions. (C) Model structure of a chromatophore photosynthetic membrane. (adapted from Sener et al. *eLife* 2016;5;e09541[92]).

In addition to LHCI, LHCII is also attributable to energy delivery to PSI[93], especially under high intensity illumination which causes phosphorylation of Lhcb3 in a LHCII trimer (**Fig. 1.6A**), control of which is considered to be the consequence of a change in mediator pool redox state.[94] The dynamic shift of LHCII between PSII and PSI serves to balance energy flux into the two photosystems, which sustains the whole photosynthetic reaction at a relatively high efficiency.

1.7.3. Rba. sphaeroides chromatophores

Long range membrane organization is also seen in the case of bacterial photosystems and is crucial for efficient photosynthetic reactions and photoprotection. The photosynthetic membrane in WT strains of *Rba. sphaeroides* is invaginated and packed with photosynthetic components. When disrupted during cell breakage, the invaginated membrane can pinch-off to form closed proteoliposomes with the diameter of ~60 nm called chromatophores.[95] RC-LH1 complexes are surrounded by pool of LH2 with the ratio of LH2 to RC-LH1 complexes varying with received radiation (**Fig. 1.6B**).[96] A representation of the structure and organization of a chromatophores is shown in **Fig. 1.6C**.[92] Not only are the photosynthetic components organized for efficient energy transfer, but also close positioning of the cyt bc_1 complex to the RC-LH1 enables fast exchange of quinone. The ATPase responsible for the production of ATP is located at LH2-riched region presumably because there is no selection pressure for close position of ATPase to photosynthetic core machinery as propagation of proton motive force is rapid.[97]

1.7.4. Other LH structures - phycobilisomes and chlorosomes

In cyanobacteria, a water-soluble antenna system is associated with its photosystem with a specific structural arrangement. The major component is called the phycobilisome (PBS) which achieves light harvesting through covalently attached open chain tetrapyrroles called bilins[98]. Depending on the chosen pigments, the absorption and emission profile of PBS will vary. Again, similar to the function of LHCII, light captured by the PBS can be efficiently distributed to both PSI and PSII, a process which relies on the mobility of the PBS, and the allocation of the PBS to PSI or PSII is environment dependent[99]. There are two states of PBS organisation, in which the PBS preferentially associates with PSII in state 1 but provides more absorption cross-section to PSI in state 2. The switch of state was considered to be controlled by an electron mediator that operates between the two photosystems (e.g. plastoquinone)[99].

Another form of antenna called the chlorosome is found attached to the photosynthetic membrane in some anoxygenic photosynthetic bacteria[100]. They are large ellipsoids of various sizes that contain stacked BChls with a limited amount of protein modules on the surface. Donation of energy harvested by chlorosomes to the intramembrane RC required protein called FMO (Fenna-Matthews-Olson protein) that is associated with the RC[101].

1.7.5. Total energy conversion and efficiency

1.7.5.1. 100 TW of energy conversion (high)

It is estimated that solar energy is converted by natural photosynthesis at rate of nearly 100 TW. This is over five times more than mankind's current global energy demand, and it is clear how important photosynthesis is to humanity in terms of food and energy. With the increasing CO_2 emission and the associated climate change, photosynthesis, either by itself or through inspiring promising techniques, provides us with a valuable approach to solving possible future energy and environmental crises. The delicate organization of photosynthetic membranes allows fast and efficient energy transfer to the RC and the quantum yield of charge separation in RCs is nearly 100%. This latter feature in particular has encouraged many attempts to directly use RCs in bio-photoelectrochemical cells for solar energy conversion, with a particular focus on PSI and purple bacterial RC and RC-LH1 complexes.

1.7.5.2. Biohybrid photoelectrochemical cells

As the key step of energy conversion by RCs is a highly quantum-efficient separation of electrical charge between opposite poles of the protein, there has been interest in incorporating them into biohybrid photoelectrochemical cells for the generation of photocurrents and photovoltages. In addition to solar to electrical power conversion[91,102–110], such devices can be applied to biosensing[111–113], photosensing[114], touch sensing[111], molecular electronics[105,115], charge storage[116] and solar fuel synthesis[117–120]. Studies have focused in the main on Photosystem I from cyanobacteria[43,121,122] and the RC and RC-LH1 complexes from purple photosynthetic bacteria, mainly *Rba. sphaeroides* [38,51,52,57]. PSII, which generates a very high redox potential of over +1 V in order to oxidize water to O_2 in the OEC[123], has been used to a much more limited extent because of its intrinsic instability due to its capacity to generate reactive oxygen species[124]. In the best case PSII can only sustain around 4000 turnovers per complex, in comparison with over one million turnovers of bacterial RCs in photoelectrochemical cell under standard 1-sun illumination[91]. Research into PSII on electrodes has principally concerned solar fuel synthesis (e.g. see Sokol *et al.*[120]).

1.7.5.3. Energy and quantum efficiencies

Despite the vast amount of energy fixed by photosynthesis the overall efficiency, defined as to stable sugar products, is only about 0.1% across the year due to many reasons[125]. Although crop plants and algae possess a higher solar energy conversion efficiency, the maximum of efficiency is only a few percent, and cannot be maintained across the year due to weather and availability of light[125]. Only a small proportion of the 100 TW of photosynthetic energy conversion can be directly accessed by us, with most energy stored in forms such as wood fibre which are hard to break down. In comparison to a theoretical limit of 32% efficiency for energy conversion by conventional silicon photovoltaic cells, green plant photosynthesis possesses a record of 12% efficiency of conversion of the energy of sunlight to that of sugar across the whole light spectrum despite an initial quantum efficiency for charge separation of nearly 100%[126]. Improving the fundamental efficiency of photosynthesis beyond this threshold and increasing our utilization of photosynthetic products are therefore attractive to both academia and industry because it provides a potential solution to global food demand and energy requirement under current critical circumstance.

1.7.5.4. Photosynthesis is not optimized for productivity

The selection pressure for the evolution of photosynthesis was not to optimize energy conversion efficiencies but rather to achieve the survival and reproduction of living organisms[127]. In fact, the appearance of photosynthesis may have been a “side-product” of the development of molecular systems for photoprotection that endowed a massive evolutionary advantage to primitive organisms exposed to high levels of ultraviolet radiation before the development of the ozone layer. The initial chlorophyll or pigment-protein complexes in membranes at the periphery of a cell may have provided some protection to internal vulnerable bio-materials such as proteins, DNA and RNA from damage, perhaps by dissipating the energy to appropriate acceptors by oxidizing the most abundant available reagent - water. The production of O_2 from water oxidation, in turn, formed the ozone (O_3) layer, strongly shielding UV light and leading to an expansion of life on the surface of the Earth. As a consequence of natural selection toward the survival and reproduction of an organism, the productivity of photosynthesis was not optimized as it was not directly subjected to selection pressure.

1.7.5.5. Nature’s limitations and strategies for improvement

The relative inefficiency of natural photosynthesis is due to several reasons including the energy losses in sequential energy and electron transfer reactions, acclimation to high light

intensity, side reactions such as photorespiration and initial incomplete spectral coverage.[127] Successful cases of improving energy efficiency have been reported by research groups targeting these aspects. For example, overexpression of the key regulators of photorespiration recovery (proteins VDE, ZEP and PsbS) has increased the efficiency of carbon fixation to 15 % in a range of model plants[128]. This was achieved by accelerating the recovery of the plant from photorespiration and so minimizing the dissipation of absorbed photon energy through the non-photochemical quenching (NPQ) cycle[128]. A study from Ort *et al.* has found that a bypass of photorespiration with an external enzyme to reduce the waste of side reactions with O_2 by Rubisco can also boost the yield of a host plant by nearly 40% in green house condition.[129] However, similar photorespiration bypass approach in rice, though improved overall photosynthesis up to 22%, resulted mixing outcome of plant yields.[130]

As outlined above, a particularly striking feature of natural photosynthesis is that different organisms exploit different parts of the available solar spectrum depending on their pigment profile. Plants exploit most of the visible spectrum between 400 nm and 700 nm, which accounts for around 50% of the solar energy arriving on the earth's surface [125], while the purple bacterial photosynthetic system accesses the near infrared region of the spectrum between 700 nm and 1000 nm[91]. Interestingly the absorbance of purple bacterial photosynthetic complexes is complementary to that of green photosynthetic organisms, which might be a consequence of evolution to avoid niche overlap[34]. In particular, purple bacteria have very little absorbance around 650-700 nm which are the red wavelengths strongly absorbed by chlorophyll in green plants, algae and cyanobacteria. To address incomplete spectral coverage in studies of solar energy conversion *in vitro* by natural pigment proteins, click chemistry has been used to add synthetic chemical chromophores with complementary spectral absorbance onto LHCs and RCs to increase the light capture at the weakly absorbed wavelengths[131,132] In the main this has involved the use of synthetic dyes which are relatively expensive and have relatively low photostabilties. In an alternative approach, Grayson *et al*. have attached yellow fluorescent protein (YFP) to the *Rba. sphaeroides* RC and demonstrated better utilisation of light by the mutant system at the wavelengths at which YFP absorbs[133].

1.8. Toward a bottom-up redesign of light capture in photosynthesis using a synthetic biology approach

The concept underlying the project described in this thesis was to address the initial spectral utility of photosynthesis by combining light harvesting components from Chl-based

photosynthesis with a RC from BChl-based photosynthesis to form a polychromatic, functional photosystems that show expanded solar energy harvesting. The motivation was to provide improved materials for testing in biohybrid devices for solar energy conversion and related technologies such as sensing. The Jones group, in work with collaborators, has explored the use of *Rba. sphaeroides* RCs and RC-LH1 complexes in a wide variety of such devices [91,134,143,135–142]. However, in general their response to visible light, especially at red wavelengths, is relatively poor. This is a limitation that can potentially be addressed by building new self-assembling photosystems, *in vitro*, that display enhanced light harvesting in regions where BChl absorbs poorly.

The component pigment-proteins used in the project were the *Rba. sphaeroides* RC and the LHCII and LHCI pigment-proteins from *Arabidopsis thaliana* (**Fig. 1.3C** and **1.5A,B**). LHCII has its main, lowest-energy chlorophyll absorbance band between 600 nm and 700 nm where the absorbance of the RC is weak (**Fig. 1.7A**) and its emission spectrum extends to 800 nm (**Fig. 1.7A,** green line), partially overlapping the RC Q_y absorbance bands. This opens the prospect that LHCII could act as an antenna for RC photochemistry and therefore produce a hybrid photosystem with a stronger response right across the nearUV-visible-nearIR range. While largely the same absorption profile exists in the case of LHCI it gives a higher far-red emission (**Fig. 1.7C**), and so could be a better energy donor to bacterial RC. As outlined above, to explore the concept of assembling photosystem hybrids a variety of genetically-modified RCs were purified from strains of *Rba. sphaeroides* lacking light harvesting complexes [83,85,91], whereas to enable genetic modification of LHCII or LHCI adapted genes were expressed in *E. coli* and the mature pigment-proteins refolded using the resulting apoproteins and purified pigments.[144,145]

Two approaches were used for the self-assembly of component proteins into functional hybrid photosystems. The first was to assemble LHCIIs and RCs into conjugates with cadmium telluride (CdTe) quantum dots (QDs), taking advantage of the polyhistidine tags engineered onto the proteins to provide strong and predictable binding to the QD surface. The QD provides a synthetic interface between the two types of protein, enabling direct energy transfer between the two as well as (because of its optical properties) mediating indirect energy transfer. The second approach was to challenge the natural system with synthetic biology from the bottom up, demonstrating the design and modularity principles inside photosynthesis. The concept explored was to adapt RCs, LHCIIs and LHCIs with additional protein components that provide an interface to facilitate self-assembly of all-protein “chimeras” (**Fig. 1.7C**). Possible genetically-encodable interfaces that were evaluated were *de novo* designed coiled coils[146], the SpyTag/SpyCatcher system[147], split inteins and a newly-designed

hybrid “Spy-Coil” system that is hybrid of a coiled coil[148] and the SpyTag/KTag/SpyLigase system[149] (**Fig. 1.7C**).

The following sections outline the properties of each of these potential synthetic and peptidyl pigment-protein interfaces.

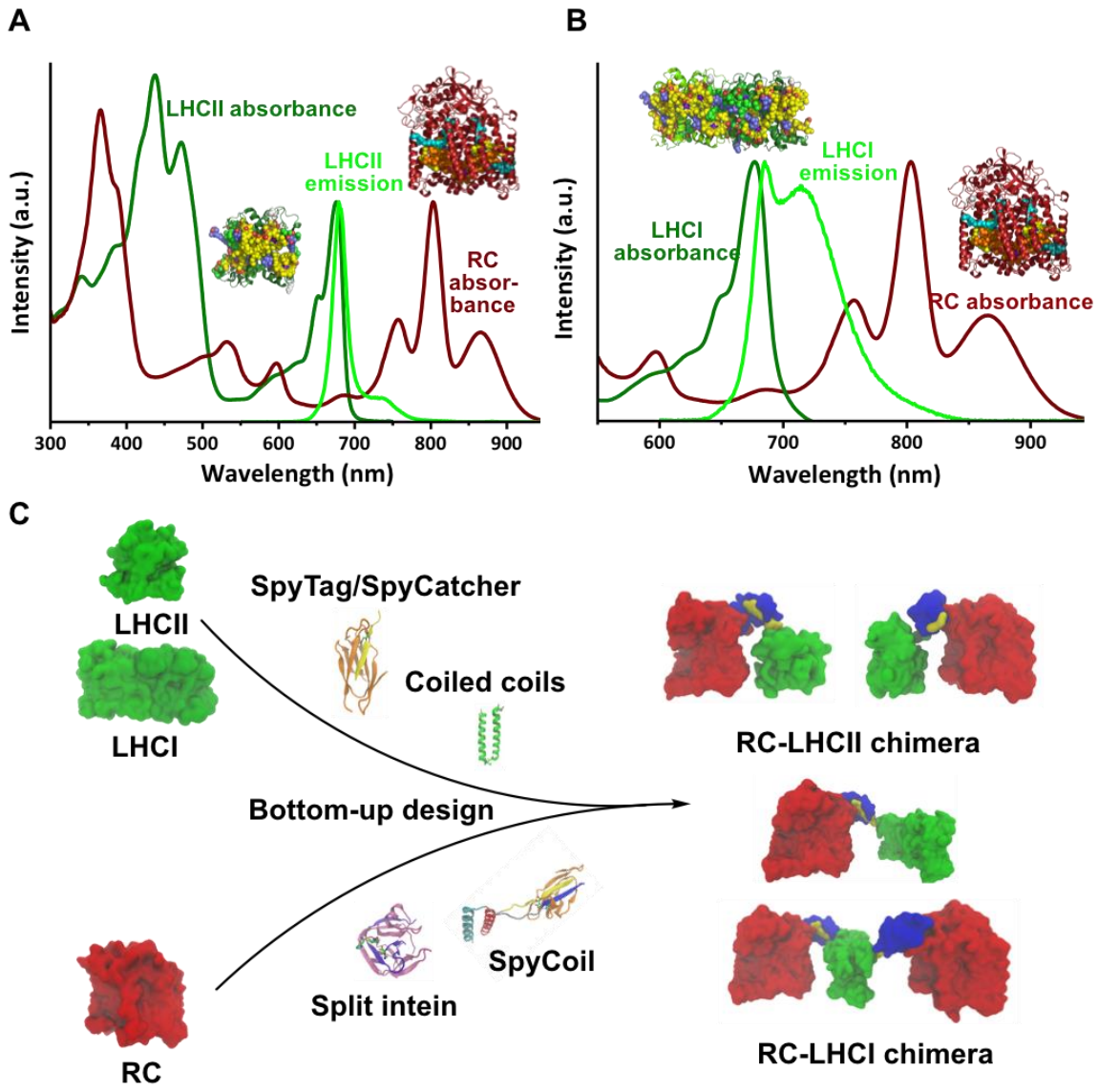


Figure 1.7 Bottom-up construction of a hybrid Chl/BChl photosystem. (A) Absorbance spectra of recombinant *A. thaliana* LHCII and the *Rba. sphaeroides* RC with the emission spectrum of LHCII, revealing the complementary absorbance profiles between LHCII and RC and spectral overlap of RC absorbance with LHCII emission. (B) Absorbance spectra of recombinant *A. thaliana* LHCI and the *Rba. sphaeroides* RC with the emission spectrum of LHCI, showing the enhanced overlap of LHCI emission with RC absorbance. (C) With the toolbox of peptidyl interfaces provided by synthetic biology, plant LHCII or LHCI can be

envisaged to self-assemble with bacterial RCs in various combinations of proteins, protein-protein orientations and oligomeric states. The examples of chimeric complexes shown were constructed with SpyTag/SpyCatcher (see Chapter 5 for details).

1.8.1. QDs – artificial light harvesting units

QDs are nanocrystals whose size is smaller than the de Broglie wavelength of the electrons so that the boundary of electron density is constrained in three dimensions[150]. The result of such "quantum confinement" is the appearance of a quantized band gap and photoluminescence properties. The smaller the nanocrystal is, the larger the energy band gap will be, and hence more blue-shifted the emission output. This tunability of QD fluorescence dependent on the size of nanocrystal gives it the potential to be utilized as an additional antenna component for a photosynthetic system, enhancing the light absorption of natural pigments in a designed way depending on the absorbance spectrum of the photosystem[151] (**Fig. 1.8A**). Once coated with amphiphilic polymers, QDs are stable in aqueous buffer and the surface property of the QD can be readily adjusted by modifying the choice of polymer forstabilization. The tuneable optical properties, high photostability and capability of undergoing multiple excitations of these semiconductor nanocrystals have been exploited in a variety of technologies including solar cells and diverse biological applications.[152–154]

1.8.2. The QD/His-tag interface

A common choice for coating QDs is to use a bifunctional ligand with a carboxyl group at its water-exposed terminus, endowing a strong negative charge to the QDs and preventing aggregation in aqueous solution. Many researchers have reported general electrostatic interactions between negatively charged QDs and biological materials including membranes and proteins. The specificity and affinity of such electrostatic interactions will depend on the QD surface chemistry and the surface charge profile of a protein of interest. The manipulation of the thermodynamics of such non-specific interactions between QDs and proteins would involve a large effort and is a case by case problem considering the vast diversity of proteins and their surface chemistry. As an alternative, conjugation of proteins with some types of QD can also be achieved with a polyhistidine tag via coordination between the histidine side chains and the surface transition metal of the QD. The QDs used in this work were composed of cadmium (Cd) and telluride (Te), of which the Cd has a strong affinity to the imidazole group

of histidine (**Fig. 1.8B**). As described in Chapters 3 and 4 in the cases of His-tagged RCs and LHCIIs the interaction with CdTe QDs was solely based on coordination to the surface Cd with no evidence of binding through electrostatic interactions. However, RC-LH1 or LH1 complexes were found to associate with CdTe QDs even without a His-tag, showing that the mechanism of protein/QD association varied as the combinatorial result of the surface properties of the QD and the particular protein used.

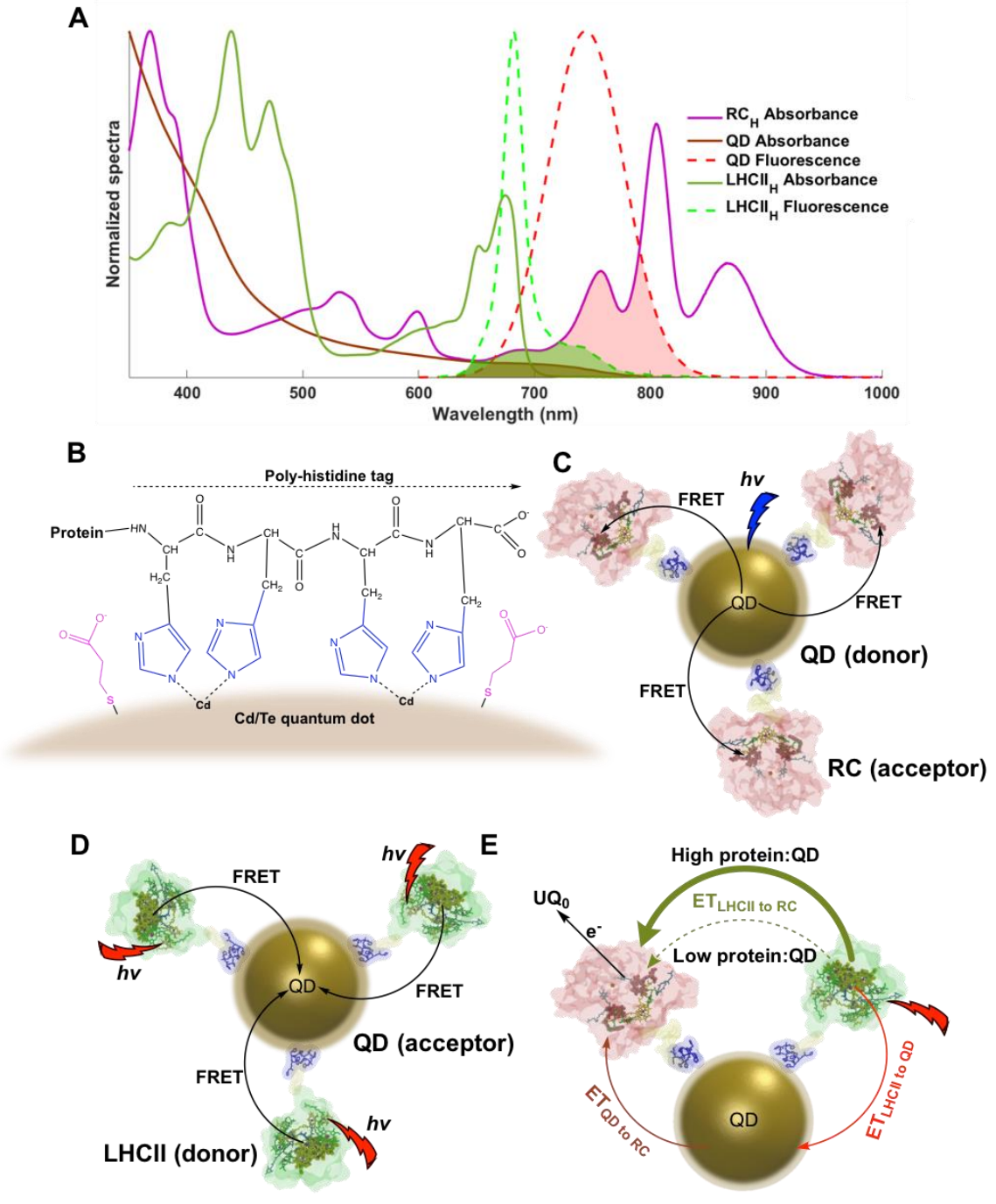


Figure 1.8 QDs filling the gap. (A) Absorbance spectra of QDs, *A. thaliana* LHCIIs and *Rba. sphaeroides* RCs overlaid with the emission spectra of QDs and LHCII. The shadowed areas highlight the overlap between RC absorbance with QD fluorescence (red), RC absorbance with LHCII fluorescence (green) and QD absorbance with LHCII fluorescence (brown). (B) Molecular mechanism of binding between the polyhistidine tag and the QD. Side chains of histidine can connect to surface cadmium by coordination bonds despite the QD being coated with an MPA (3-mercaptopropionic acid) bifunctional ligand (purple). (C) Possible energy coupling of RCs (acceptor) with a QD (donor) via FRET in a RC/QD conjugate. Plausible pigments responsible for energy transfer are highlighted with spheres. (D) Possible energy flow between LHCIIs (donors) and a QD (acceptor) via FRET in a LHCII/QD conjugate. Lowest energy pigments presumably used for energy transfer are highlighted with spheres. (E) Possible flows of energy in a LHCII/RC/QD tri-component conjugate. The photon energy harvested by LHCII can be delivered to the RC either by direct energy transfer ($ET_{LHCIItoRC}$) or indirectly through the QD ($ET_{LHCIItoQD}$ and ET_{QDtoRC}).

1.8.3. Tuning QDs for photoprotein energy coupling

In addition to providing a hub for the assembly of LHCIIs and RCs into a single photosystem, by selecting QDs with suitable properties (6.5 nm diameter, 750 nm emission) they also could potentially bridge the energy gap between LHCII and the RC by being able to act as both an energy acceptor and donor (**Fig. 1.8A**)[155]. The absorbance of the QDs extended from 800 nm into the UV, hence overlapping with the emission spectrum of LHCII between 650 nm and 800 nm (**Fig. 1.8A**). In LHCII/QD conjugates this could facilitate LHCII→QD energy transfer (**Fig. 1.8D**). The emission spectrum of the QDs covers 650 nm to 850 nm, overlapping mainly with the Q_y absorbance bands of the BPhes and monomeric BChls of the RC (**Fig. 1.8A**). In RC/QD conjugates this could facilitate QD→RC energy transfer (**Fig. 1.8C**). In a tri-component conjugate (**Fig. 1.8E**), excitation of LHCII could initiate the delivery of energy from the Chl light harvesting system of LHCIIs to the BChl charge separation system of the RCs either directly or via the energy conduit provided by the QD assembly hub.

1.8.4. The arsenal of membrane protein compatible linkers

Synthetic biology has created a range of exciting peptidyl-based linking strategies that can operate without external inputs, enabling self-assembly of multicomponent hybrid complexes. Three such systems were evaluated in this work, plus a fourth, hybrid system that was designed as described in Chapter 6.

1.8.4.1. SpyTag/SpyCatcher

The SpyTag/SpyCatcher system[147] was derived from the CnaB2 adhesin domain expressed by *Streptococcus pyogenes.* It is a subunit of the FbaB extracellular protein using for binding to fibronectin. An internal isopeptide bond spontaneously forms within the CnaB2 domain between aspartic acid and lysine side chains, enhancing the stability of the protein. A protein fusion system was developed by splitting the CnaB2 domain into two parts, each containing one of the two reactive side-chains. By engineering the association surface between the two parts, efficient ligation was achieved, and the optimized pair was named as SpyTag and SpyCatcher[147] (**Fig. 1.9A**). SpyTag is a short strand containing the reactive aspartic acid, while SpyCatcher is the rest of the CnaB2 domain with the reactive lysine. Upon formation of SpyTag/SpyCatcher complex, the side chains of the reactive Lys and Asp are linked by a covalent bond (**Fig. 1.9A**).

Once genetically incorporated with proteins of interest, SpyTag/SpyCatcher will trigger the self-assembly of two modified proteins into an ultra-stable chimera product in a reaction that tolerates a wide range of buffer conditions. SpyTag or SpyCatcher can be engineered at either terminus of a protein chain and can even be inserted inside a host protein sequence as long as it does not disrupt protein folding. Based on the crystal structure of the SpyTag/SpyCatcher complex, the size of the interacting components could be reduced by removing redundant sequences that no longer contribute to the interaction, and these optimised versions were used in this project.[156] Although SpyTag/SpyCatcher created a sizeable adduct between the two proteins of interest that may have limited their separation distance, its high buffer condition tolerance, self-driven reaction without side products and the high efficiency and ultrastability of the ligation made it the most successful of the approaches tested for building photosynthetic supracomplexes with high precision. The recent development of another pair of split proteins called SnoopTag/SnoopCatcher[157] provides a toolbox of orthogonal linkers for the construction of more complicated protein assemblies with multiple components as building blocks.

1.8.4.2. *De novo* designed coiled coils

The coiled coil is the most common tertiary protein fold found in nature[158]. Coiled coils are formed by α-helices that can cluster in various oligomers, orientations and combinations. Through study of coiled-coils for many years, they are now one of the most understood protein structures, whose specificity and affinity can be tuned accurately for desired properties[159]. Investigation of dimeric coiled-coils revealed a repeated residue pattern with the heptad unit, *HPPHPPP* (H-hydrophobic, P-polar), and the repeat sequence can be annotated by positions ***abcdefg***. As the result of repeated occupation of hydrophobic residues at the ***a*** and ***d*** positions, a hydrophobic seam is present on one side of the helix, which leads to dimerization of two helices by hydrophobic interactions (**Fig. 1.9B**, silver side chains in top-down view)[148].

However, the hydrophobic residues at the ***ad*** positions are not the only determinants of the coiled-coil structure. In the case of the heterodimeric coiled coil using in this project, the CCA/CCB pair, the two helices carry oppositely charged amino acids (Asp – CCA; Lys – CCB)

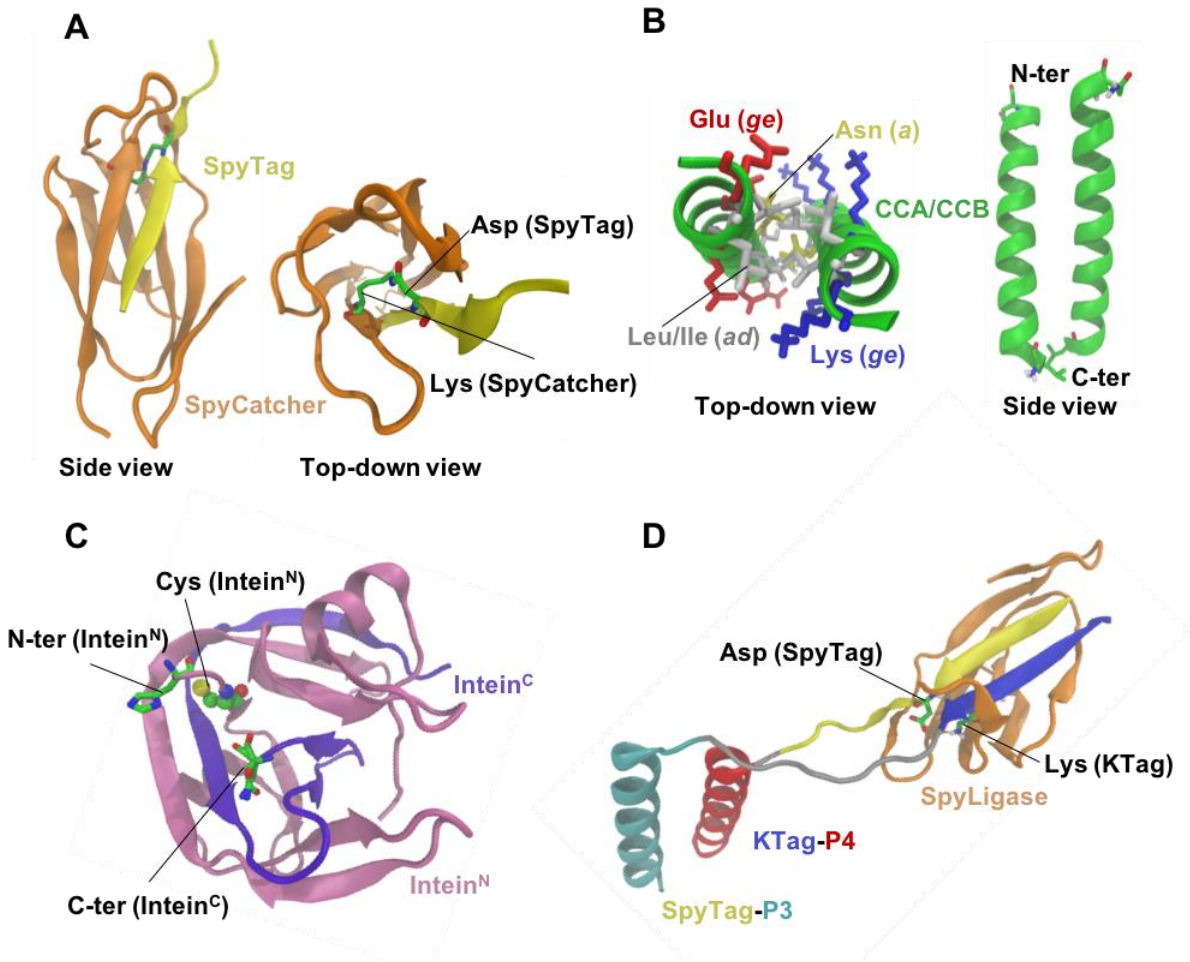


Figure 1.9 Peptidyl interfaces for chimera self-assembly. (A) Complex of SpyTag (yellow) and SpyCatcher (orange). Side chains of reactive Asp on SpyTag and Lys on SpyCatcher are highlighted and an isopeptide bond can be observed from the top-down view. Model generated from PDB entry 4MLI[156]. (B) *De novo* designed coiled coil (CCA/CCB pair). The backbone of the coiled coil is shown as a green ribbon. The hydrophobic core at the *ad* positions (silver) provides affinity while counter charged amino acids (Glu - red; Lys - blue) at the ***ge*** positions provide specificity for parallel heterodimeric packing. The Leu at ***a*** and Ile at ***d*** define the oligomeric state to be a dimer and an Asn pair at ***a*** is also required for dimerization. Model generated from ISAMBARD[160]. (C) Complex of the split intein (PDB entry Npu4LX3[161]). A transpeptidyl reaction will take place at the N-terminus (N-ter) of the intein N-ter domain ($Intein^N$) and the C-terminus (C-ter) of the intein C-ter domain ($Intein^C$), which leads to cleavage of the intein and joining of the two parts attached to $intein^N$ and $intein^C$ with the assistance from catalytic Cys at both the N-ter of $Intein^N$ and the C-ter of $Intein^C$ (only one resolved in the crystal structure). (D) SpyCoil with SpyLigase. A heterodimeric coiled coil (P3 – cyan; P4 - red) connected to SpyTag (yellow) and Ktag (blue), respectively, via flexible linkers (silver). In the presence of SpyLigase (orange), an isopeptide bond will be formed between the Asp on SpyTag and Lys on KTag. Design of SpyCoil is discussed in Chapter 6.

---

at the ***e*** and ***g*** positions so that Coulombic interactions between the ***eg*** residues will encourage parallel heterodimeric packing[162] (**Fig. 1.9B**, blue/red side chains in top-down view). Asparagine (Asn) at position ***a*** is also a crucial design principle for dimerization and contamination of other oligomeric states will appear if without Asn (**Fig. 1.9B**, yellow side chains)[163]. Some examples of adaption of CCA/CCB to interface modules of interests have been shown[164,165] but the exact configuration of produced complexes was rarely studied and is poorly understood.

1.8.4.3. The split intein

Inteins are a family of proteins that are found widely spread in most life forms with little contribution to the host organisms' survival. Inteins are sequences that can excise themselves from a precursor polypeptide and join the flanking sequences (exteins) together with a polypeptide bond. If an intein domain is split into two parts, either through design or natural occurrence, they can be used for post-translational modification when attached to host proteins of interest. Because, once folded, the split intein domain cleaves itself from the

attached sequences, this leads to the creation of a “clean” linking system in which the N-terminal intein domain on one host protein and the C-terminal domain on a second host protein will combine and self-cleave to leave a short linking “scar” that consists of three amino acids. In the case of the designed split intein system Cfa used in this work, the scar sequence is “CFN”[166]. The Cfa split intein system is based on a consensus sequence of natural split inteins from a database and the crystal structure of one of the most studied split inteins (Npu) was used to generate the model in **Fig. 1.9C** (PDB entry 4LX3[161]). Two key reactive cysteines, one on the Cfa^{N}-domain and another on the Cfa^{C}-domain, are crucial to the ligation reaction. Once in the reduced form and correctly steered, cysteine substitute the peptide bond forming a cyclic reactive species and then transfer the attached amine group to the adjacent free carboxyl terminal of Cfa^{C}-domain, generating a new peptide bond linking two host proteins together. Although in principle this is a very effective way of linking two proteins using a component that is not retained in the final product, intein splicing can be affected by side reactions leading to low efficiency[167]. Abortion of the desired splicing reaction is caused by hydrolysis of such reactive species, resulting in detachment of the intein domain without peptidyl-linkage being formed between the two host proteins. The abortive reactions seemed to be more preferred with the increasing in the size of the protein they attached to[167].

1.8.4.4. SpyCoil – combining the merits of coiled coils and the Spy-system

The fourth linking system evaluated in this work was a newly-designed hybrid system named SpyCoil, which combined the SpyTag/KTag/SpyLigase system with a heterodimeric coiled coil.

The SpyTag/KTag/SpyLigase protein fusion system is the result of further splitting the SpyCatcher protein into two parts. The reactive lysine is included in the KTag peptide and the rest of SpyCatcher is renamed as SpyLigase[149]. The concept is that SpyTag and KTag can be engineered onto proteins of interest and, upon mixing the two tags and SpyLigase, two modified proteins can be covalently linked when the overall Spy-complex is correctly formed. However, due to the over-splitting of CnaB2 domain, this tri-component Spy-system is only functional in the presence of a high concentration of trimethylamine N-oxide (TMAO). This chemical is commonly used for stabilizing protein structure, which indicated that the SpyTag/KTag/SpyLigase combination seemed not to be able to form a stable complex in normal aqueous solution. The idea behind the SpyCoil design was therefore that if one can provide additional affinity to pre-assemble part of the complex, the activity of the over-split tri-component Spy-complex might be restored. Moreover, because the activity depends on the extra affinity provided by the additional binding motif, the specificity can be readily tuned by

using designed units. *De novo* designed heterodimeric coiled coils were perfect candidates to fulfil such a purpose since they are well understood, and many examples of designed coiled coils have been reported previously. In this study, a heterodimeric pair called P3 and P4 from Bromley *et al* were used among other two orthogonal pairs of coiled-coil sequences[148].

1.9. Aims of this study

The aims of the research described in this thesis were:

(1) To explore the use of QDs as hubs for the self-assembly of RC/QD, LHCII/QD and RC/LHCII/QD conjugates.
(2) To characterise the composition, thermodynamics of assembly and energy transfer properties of RC/QD, LHCII/QD and RC/LHCII/QD conjugates.
(3) To explore the use of programmable peptide interfaces for the self-assembly of LHCII-RC and LHCI-RC chimeras.
(4) To characterise the composition and energy transfer properties of LHCII-RC and LHCI-RC chimeras.

In Chapter 2 the materials and methods used in this work are described.

In Chapter 3 the self-assembly of RC/QD conjugates is described. Their architecture is explored, the thermodynamics of their assembly are characterised and the mechanism of energy transfer from QDs to attached RCs is identified using RCs with mutations that alter their pigment content.

In Chapter 4 the self-assembly of LHCII/QD and RC/LHCII/QD conjugates is described. Their compositions and architectures are determined and the thermodynamics of their self-assembly are characterised. By comparison of the di-component and tri-component conjugates the contributions of direct LHCII→RC and indirect LHCII→QD→RC energy transfer in the tri-component system are disentangled and the efficiencies of the parallel energy transfer routes estimated.

In Chapter 5 the self-assembly of LHCII-RC and LHCI-RC chimeras using the SpyCatcher/SpyTag system is described. Evidence for LHC→RC energy transfer is presented and the efficiency of energy transfer in chimeras of different compositions estimated. The

ability of chimeras to generate photocurrents on nanostructured silver electrodes is also explored.

In Chapter 6 the utility of coiled coils and a split intein system for integral membrane pigment-protein fusion is evaluated. Design and implementation of the SpyCoil protein fusion system is also described, and its utility explored using test proteins including LHCII.

Chapter 7 presents a brief summary of the findings from this work and a consideration of future directions this work could go in.

Chapter 2

Materials and Methods

2.1. Bacterial strains

2.1.1. Escherichia coli strains

The following strains of *E. coli* were used.

Table 2.1. *E. coli* strains used in this work.

Strain	Features	Ref.
DH5α	DNAase deficient mutant facilitating plasmid transformation.	Bethesda[168]
S17-1	λpir integrated mutant for plasmid transfer to *Rba. sphaeroides*.	Simon el al.[169]
BL21(DE3)	Protease deficient strain for T7 RNA polymerase expression of proteins under IPTG-inducible lac UV5 promoter.	Studier et al.[170]
NiCo21(DE3)	Derivative of BL21(DE3) engineered for reduced endogenous binding to IMAC resins.	Robichon et al.[171]
C43(DE3)	Derivative of BL21(DE3) with resistance to expression of toxic proteins.	Miroux et al.[172]
C43(DE3) pLysS	Derivative of C43(DE3) expressing T7 lysozyme to reduce background expression of T7 RNA polymerase.	Wagner et al.[173]
Rosetta™ 2	Derivative of BL21 with 7 tRNAs for rare codons and suitable for eukaryotic protein expression	Novy et al.[174]
NEB® 5-alpha Competent	Derivative of DH5α used for high-efficiency transformation	NEB

2.1.2. Rhodobacter sphaeroides strains

The following strains of *Rba. sphaeroides* were used.

Table 2.2. *Rba. sphaeroides* strains used in this work.

Strain	Features	Ref.
DBCΩ	Replacement of LH2 genes in *puc* operon; $Strp^R$	Jones et al.[86]
DD13	Replacement of LH1, RC and PufX genes in *puf* operon, and LH2 genes in *puc* operon; $Strp^R$, Neo^R	Same above
DD13/G1	Spontaneous mutant of DD13 with neurosporene (green) as principal carotenoid	Same above

$Strp^R$ = streptomycin resistant; Neo^R – neomycin resistant

2.2. Bacterial growth, harvesting and storage

2.2.1. Rhodobacter sphaeroides

Culture of *Rba. sphaeroides* was carried out using M22+ medium[175], which is an enhanced form of M22 medium[176] (see **Appendix 2.1**, Section 2.20). Liquid cultures were supplemented with 0.1 % (w/v) tryptone before autoclaving. Preparation of solid media was accomplished by adding 1.5 % (w/v) agar to liquid media and then autoclaving. A filter sterilised vitamins solution (see **Appendix 2.2**, Section 2.20) was added to solid media after cooling to ~50°C, immediately before plates were poured, and to cooled liquid media immediately before inoculation. Filter sterilised antibiotic solutions were also added to cooled culture media as required. Concentrations were: tetracycline (Tet) – 1 μg mL^{-1}; neomycin (Neo) – 20 μg mL^{-1}; streptomycin (Strp) – 5 μg mL^{-1}.

Plates of inoculated M22+ were incubated at 34°C in the dark until colonies grew (typically 5-7 days). Growth of liquid cultures was carried out under dark/semi-aerobic conditions. First a starting culture of 70 mL M22+ in a 100 mL conical flask was inoculated using cells scraped from a glycerol stock (see below) and cultured overnight under 34°C at 180 revolutions per minute (rpm) in a darkened orbital incubator. Then this starting culture was transferred to 1.5 L of M22+ in a 2 L conical flask and further grown for 36 to 48 h at 34°C and 180 rpm until the cell density reached stationary state. Larger volumes were grown by subculture as required (typically 12-18 L of culture volume for protein purification). Cells from large cultures were harvested using a Sorvall RC3BP+ swing-out centrifuge at 4,000 rpm and 4°C for 20 min. Cell pellets were stored at -20°C before use.

Stocks of viable cells for inoculation were kept at -80°C as concentrated suspensions in a filter sterilised 30% glycerol/70% Luria-Bertani (LB) broth that was snap-frozen to improve cell viability.

2.2.2. Escherichia coli

E. coli were grown in LB broth (see **Appendix 2.3**, Section 2.20). For colony culture, LB agar plates were prepared by adding 1.5% (w/v) agar to liquid media and autoclaving. Appropriate antibiotics were mixed with warm LB/agar medium if required. Inoculation of *E. coli* was achieved by spreading around a 100 µL volume of cells and spread plates were incubated overnight at 37°C. Small scale liquid culturing of *E. coli* was accomplished in 10 mL volumes of LB broth in 30 mL universal bottles. Large scale liquid culturing of *E. coli* was accomplished in 1 L LB medium in a 2 L conical flask with appropriate antibiotics, and growth was at 37°C at 250 rpm in an orbital incubator. OD_{600} or OD_{550} was used to indicate the cell density. Antibiotic final concentrations were: Tet - 50 µg mL^{-1}; ampicillin (Amp) – 200 µg mL^{-1}; kanamycin (Kan) - 50 µg mL^{-1}. Glycerol stocks of viable cells for inoculation were kept at -80°C in the same way as for *Rba. sphaeroides* (see above).

2.3. Expression vectors

The following expression vectors were used.

Table 2.3. Expression plasmids.

Plasmid	Features	Ref.
pUCXB-1	For modification of RC *pufLM* genes; Amp^R	McAuley-Hecht et al.[177]
pRKEH10D	For mobilisation and expression of RC *pufLM* genes; Tet^R	Jones et al. [86]
pOPINE	T7/LacO for IPTG induction in *E. coli* ; Amp^R	Bercow et al.[178]
pBAD/Myc-His	araBAD for arabinose induction in *E. coli*; Amp^R	Invitrogen
pET28a	T7/LacO for IPTG induction in *E. coli*; Kan^R	Novagen

Amp^R – ampicillin resistance; Tet^R – tetracycline resistance; Kan^R – kanamycin resistance.

2.4. Molecular Biology

2.4.1. Preparation of plasmid DNA

Plasmids were prepared from 10 mL *E. coli* overnight cultures using a miniprep kit (QIAprep® Spin, QIAGEN). Purified plasmid DNA was quantified with a NanoDrop spectrophotometer (Thermo Fisher) and concentration was determined from the absorbance spectrum. Plasmids were stored at -20^{o}C until use.

2.4.2. Preparation of PCR products and digested linear DNA

Polymerase chain reaction (PCR) products or restriction digested DNA fragments were firstly separated from other components by agarose gel electrophoresis and the corresponding band was excised and processed using a gel extraction kit (Qiagen™). Purified DNA fragments were examined by agarose gel electrophoresis and concentrations were estimated using the NanoDrop™ spectrophotometer. Buffers used for agarose gel electrophoresis were either TBE (Tris-borate-EDTA; Tris - tris(hydroxymethyl)aminomethane; EDTA - ethylenediaminetetraacetic acid) or TAE (Tris-acetate-EDTA) – see **Appendix 2.4**, Section 2.20. Final DNA products were stored at -20^{o}C until use.

2.4.3. Restriction digests and ligations

Restriction digests and ligations were used to shuttle DNA fragments between vectors. Plasmid DNA was incubated with the desired enzyme(s) under the correct buffer conditions for at least 3 h. The required restriction fragments were identified and purified by agarose gel electrophoresis and DNA gel extraction. Purified DNA fragments were incubated with T4 DNA ligase at 16^{o}C overnight in ligation buffer supplied with the enzyme. The molar ratio of insert to vector was kept at 5:1. Final ligated products were transformed into chemically-competent *E. coli* DH5α (**Table 2.1**) after diluting 5 times with MilliQ water.

2.4.4. PCR amplification

PCR amplification of designed DNA sequences was carried out using Q5® High-Fidelity DNA polymerase (New England Biolabs (NEB)). Annealing temperatures and extension times were determined using NEB online tools. The cycle number was usually set to 35.

2.4.5. DNA transformation into chemically competent cells

Transformation was accomplished using chemically competent *E. coli* cells. Home-made competent cells were prepared by harvesting cultures of *E. coli* (OD_{550} = 0.35-0.5) and resuspending cells in RFI buffer (see **Appendix 2.5**, Section 2.20). After incubating on ice for 30 min, the cells were centrifuged at 4 ^{o}C and resuspended in RFII buffer (see **Appendix 2.5**, Section 2.20) to endow competence. Cells were snap frozen in 200 µL aliquots and stored at -80^{o}C. Home-made competent cells were used for most of the work described with the exception that commercial competent cells were used for Q5® Site-Directed Mutagenesis to overcome a low yield of product.

Transformation was achieved by adding plasmid DNA to competent cells that had been thawed on ice. The cells were incubated in ice for 30 mins and then heat shocked at 42^{o}C for 30s. The cells were returned to ice for 5 mins and then allowed to recover in 4 x volumes of LB medium for 1 hr in an orbital incubator at 37^{o}C and 250 rpm. Finally, the cells were spread on an LB agar plate supplemented with appropriate antibiotics and the plate was incubated at 37^{o}C overnight.

2.4.6. DNA sequencing

All modified plasmids were checked using the Sanger DNA sequencing service provided by Eurofins Scientific with appropriate primers. Sequencing results were aligned with the corresponding plasmid sequence using the ApE plasmid editor (http://jorgensen.biology.utah.edu/wayned/ape/) or the SnapGene Viewer (http://www.snapgene.com/).

2.4.7. Oligonucleotide and gene design and synthesis

The design of DNA oligonucleotides was conducted according to assumption that they would generally be used for some form of PCR-based mutagenesis or assembly. Some common criteria were followed such as ensuring that the annealing temperature was in the range 60^{o}C to 72^{o}C; primers finished with C or G at the 3' end; there was no obvious homo/hetero-dimerization and off-targeting when being used in the PCR mixture and no extensive repeated sequences that were difficult to synthesise. Designed DNA oligonucleotides were synthesized by Eurofins as unmodified single stranded DNA, dissolved in MilliQ water to 100 ng mL^{-1} on receipt and stored at -20^{o}C before use.

Using online tools provided, designed genes were manufactured by either Eurofins or GeneArt synthesis (Thermo Fisher) as inserts in a supplied vector. Optimization of codon

usage was performed for expression in *Rba. sphaeroides* but not for expression in *E. coli* because the Rosetta™ 2 strain was used for most expression work. Whole genes or gene fragments were shuttled into the appropriate expression vector (e.g. pET28a, pUCXB-1) by restriction digest and ligation.

2.4.8. Site-directed mutagenesis

Point mutations were made using the QuikChange mutagenesis protocol (Stratagene). A pair of primers with the required nucleotide changes flanked by fully complementary sequences was used to amplify template plasmid DNA by PCR using *Pfu Turbo* DNA polymerase. Before transforming the PCR product, template DNA was digested by incubating with DpnI for 1 h at 37^{o}C.

Small insertions or deletions in a DNA sequence (<100 bp) were achieved using the Q5® Site-Directed Mutagenesis Kit (NEB) with designed primers. The PCR products from this were usually blunt end fragments and were joint by the KLD enzyme mixture provided in the kit. Final circularized plasmids were transformed into NEB® 5-alpha Competent *E. coli* cells (**Table 2.1**).

2.4.9. Extensive modification of sequence by Gibson assembly

Extensive modification of a gene or shuffling of the order of motifs within a gene was achieved using a Gibson Assembly® kit following the protocol from manufacturer (NEB). Of note, 20 nucleotide (nt) overlap domains with every adjacent fragment were always used. PCR products were cleaned and purified as described in Section 2.4.2. During the reaction at 50^{o}C, a 10-fold molar ratio of short fragments (<200 bp), or 5-fold molar ratio of fragments with size >500bp, was used with respect to the plasmid backbone (5 kb ~ 6kb). Final circularized plasmids were again transformed into NEB® 5-alpha Competent *E. coli* cells (NEB).

2.4.10. Horizontal gene transfer from *E. coli* S17-1 to *Rhodobacter*.

Modified RC genes were constructed initially within plasmid pUCXB-1 which encodes *pufL* and *pufM* for the RC L and M polypeptides. Modified genes were transferred to expression vector pRKEH10D (Tet^{R}) as an XbaI-BamHI restriction fragment. The modified expression vector was introduced to *E. coli* strain S17-1 (Table 2.1.) which was used for conjugative transfer of the modified genes to *Rhodobacter* strains lacking genomic copies of those genes (Table 2.2). The transfer protocol was adapted from Hunter and Turner[175]. First, transformed *E. coli* S17-1 was grown overnight on LB agar with Tet and a match-head-sized scrape of cells

was resuspend in 50 µL filter sterilized LB. An overnight 70 mL culture of the recipient *Rba. sphaeroides* strain (usually DD13 (Neo^R) – see Table 2.2) was harvested by centrifugation and the cell pellet resuspended in 2 mL of M22+. A 5 µL aliquot of *E. coli* was mixed with 100 µL of *Rba. sphaeroides* and the mixture pipetted onto a dried LB/agar plate. The plate was then cultured for 7 hrs at 34^{o}C in the dark to allow conjugative transfer. The patch of cells was then resuspended in 200 µL M22+. A ten-fold dilution and the original concentrated suspension were spread onto M22+ agar plates supplemented with Tet and Neo. The plate was incubated at 34^{o}C in the dark for approximately a week until the appearance of colonies. Selected colonies were re-streaked onto another M22 agar plate and a lump of cells scraped from this second plate were used to inoculate 70 mL of M22+ after growth under dark/semiaerobic conditions for 24 h.. Glycerol stocks were then prepared (Section 2.2.1).

2.5. Protein expression in *Rhodobacter* and protein purification

2.5.1. General procedures

Centrifugation was carried out using a range of bench top centrifuges and microfuges, and with a Sorvall RC6 centrifuge with a SS34 fixed angle rotor. Ultracentrifugation was carried out using a Beckman Optima L-100XP ultracentrifuge with 45Ti and 70Ti fixed angle rotors and a SW 32 Ti swinging bucket rotor. A Sorvall Discovery 90 ultracentrifuge was also used with a TH-641 swinging bucket rotor. Protein purification was carried out with a variety of Akta protein chromatography systems, one of which was made available by the BrisSynBio Centre for Synthetic Biology, UoB. Absorbance spectroscopy was carried out with an Agilent Technologies Cary 60 UV-Vis Absorbance Spectrophotometer or a Beckman DU640 Absorbance Spectrophotometer. Fluorescence spectroscopy was carried out with an Agilent Technologies Cary Eclipse Fluorescence Spectrophotometer kindly made available by Dr. Paul Curnow, School of Biochemistry, UoB.

2.5.2. Purification of His-tagged RCs

The protocol used was adapted from that described by Swainsbury *et al.*[113]. Cells of *Rba. sphaeroides* grown under dark/semiaerobic conditions and harvested by centrifugation were suspended in 20 mM Tris (pH 8.0), supplemented with protease inhibitor and DNAase, and lysed in a high-pressure cell disruptor (Constant Systems) at 21,000 psi (pounds per square inch). Cell debris was removed by centrifugation at 18,000 rpm for 15 mins 4^{o}C using an SS-34 rotor. The supernatant was then incubated in the dark with 1.5 % (v/v) lauryldimethylamine N-oxide (LDAO) and 200 mM NaCl at 4^{o}C for least 1 h. After ultracentrifugation at 38,000 rpm for 30 min 4^{o}C in a Ti45 rotor the supernatant was collected, and initial purification carried out

using a 20 mL Ni-NTA (nitrilotriacetic acid) column (GE Healthcare). The nickel column was equilibrated with five column volumes (CV) of RCHisA buffer (20 mM Tris (pH 8.0)/300 mM NaCl/0.1% LDAO/20 mM imidazole), the protein was loaded, the column was washed with 20 CV of RCHisA and the His-tagged RCs were eluted with a linear imidazole gradient up to 500 mM. The eluate was concentrated and then further purified using a Superdex 200 gel-filtration column in 20 mM Tris (pH 8.0)/0.1 % LDAO (Tris-LDAO). If required, buffer/detergent exchange was accomplished at the same time. The eluted fractions with the lowest 280 nm (protein) to 804 nm (RC BChl) absorbance ratios ($A_{280}/A_{803} < 1.5$) were kept and concentrated, and then aliquoted in small volumes (~100 μL). Finally, aliquots were stored at -80°C.

2.5.3. Purification of RCs lacking a His-tag

The procedure to purify RCs without a His-tag was based on that described by McAuley-Hecht *et al.*[177]. The procedure was as the same as for purification of His-tagged RCs up to membrane solubilization with Lauryldimethylamine oxide (LDAO) and ultracentrifugation to remove membrane debris. Protein in the supernatant was desalted by dialysis against Tris-LDAO buffer overnight and then bound to a diethylaminoethyl cellulose anion exchange column (DE52, Whatman) equilibrated in the same buffer. Then the column was washed with 2x CV of Tris-LDAO, 2x CV of Tris-LDAO + 50 mM NaCl and 2x CV of Tris-LDAO + 100 mM NaCl. RCs were eluted with 2x CV of Tris-LDAO + 200 mM NaCl and collected in 5 mL fractions. Fractions with A_{280}/A_{803} ratio below 2.0 to 3.0 were pooled and fractions with the same ratio between 2.0 to 6.0 or 3.0 to 8.0 were subjected to a second round of the anion exchange step, after which fractions with a A_{280}/A_{803} ratio <3.0 were pooled with saved fractions from the first step. After dialysis the RC was further purified with a Q Sepharose HP anion exchange column equilibrated in Tris-LDAO. After loading and washing with Tris-LDAO buffer, the RCs were eluted by a linear salt gradient from 0 to 500 mM in Tris-LDAO and the fractions with $A_{280}/A_{803} < 2.0$ were collected. Final purification was achieved by a step of gel-filtration chromatography using a Superdex 200 column. Detergent/buffer exchange was also accomplished at this step if required. Fractions with $A_{280}/A_{803} < 1.6$ were deemed as the final pure product that was concentrated, aliquoted, snap-frozen and stored at -80°C.

2.5.4. Purification of His-tagged RC-LH1 complexes

Purification of RC-LH1 complexes with a His-tag on the RC component was carried out according to the same protocol as used for His-tagged RCs with the exception that *n*-dodecyl β-D-maltoside (DDM) was used as the detergent throughout. A final concentration of 1 % DDM was used for protein solubilisation from membranes and 0.04 % DDM replaced 0.1 % LDAO

in all buffers. Fractions after gel-filtration were in 20 mM Tris (pH 8.0)/0.04 % DDM (Tris-DDM). Fractions with A_{873}/A_{280} in the range 2.0 – 2.3 were collected, pooled, concentrated, aliquoted, snap-frozen and stored.

2.6. LHC expression in *E. coli* and *in vitro* reconstitution

The LHCII and LHCI complexes used in this work were genetically modified. To enable this, apoproteins were expressed in *E. coli* and the holoproteins refolded using purified pigments, according to procedures previously published for LHCII [72,145,179] and LHCI[73,74,76,180].

2.6.1. LHCII expression as inclusion bodies

Designed genes encoding modified *A. thaliana* LHCII apoproteins were cloned into a pET28a vector and were expressed in *E. coli* strain Rosetta™ 2 (**Table 2.1**). Expression started with an overnight 10 mL seed LB culture supplemented with Kan. The seeding culture was scaled up to 1 L and grown in an orbital incubator at 37°C and 250 rpm to an OD_{600} = 0.6 ~ 0.8. Induction of expression was achieved by adding between 0.4 to 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 31°C to 37°C for 4 to 8 h, depending on the difficulty of expression. Cells were harvested by centrifugation and stored at -20°C before use.

Thawed cell pellets were resuspended by vortexing in lysis buffer (50 mM Tris, 2.5% (w/v) sucrose, 1 mM EDTA, pH 8.0) with protease inhibitor. Fully resuspended cells were lysed by passing through a Constant Systems cell disruptor at 25,000 psi with 10 mg lysozyme per litre of original culture. Separation of inclusion bodies (IBs) from other cell fractions was accomplished by centrifugation of the cell lysate at 12,000 g for 15 min at 4°C. Inclusion bodies (IBs) from the pellet were washed once with Detergent Buffer (20 mM Tris (pH 7.8)/200 mM NaCl/2 mM EDTA/1 % sodium deoxycholate/1 % octyl β-D-glucopyranoside (OG)/10 mM dithiothreitol (DTT)), three times with Triton Buffer (20 mM Tris (pH 8.0)/1 mM DTT/0.5 % Triton X-100) and a final wash with TE Buffer (20 mM Tris (pH 8.0)/1 mM EDTA) to remove excess detergents. Purified IBs were stored at -80°C as pellets and the protein purity was checked by sodium dodecyl sulphate – polyacrylamide gel electrophoresis (SDS-PAGE) using quick Coomassie stain staining (Section 2.10.1). The concentration of total protein was measured by a *DC*™ protein assay (BioRad). In most cases nearly all of the protein in the IBs was the desired LHCII apoprotein.

2.6.2. Pigment extraction

Total Chl and carotenoids were extracted from fresh spinach leaves (purchased from a local supermarket). Leaves were homogenized in cold Grinding Buffer (0.4 M sorbitol/10 mM NaCl/5 mM $MgCl_2$/0.5% (w/v) milk powder/0.1 M Tricine (pH 7.8)) using a Waring immersion blender (20 g leaves per 100 mL buffer). Chloroplasts were separated by filtration through nylon material (20 μm pores) and centrifugation of the filtrate at 1,500 g for 10 min at 4°C. The chloroplast pellet was washed by twice resuspending in 50 mL Wash Buffer (50 mM sorbitol/10 mM EDTA (pH 8)/5 mM Tricine (pH 7.8)) and centrifuging at 10,000 g for 10 min at 4°C.

Having removed all supernatant, total pigment was extracted in the dark by adding 20 mL of 80 % acetone buffered with sodium carbonate to the chloroplasts and incubating on ice with occasional vortexing. Unextracted material was removed by centrifugation at 12,000 g for 15 min at 4°C. The supernatant was added to a separating funnel, 0.4 volumes of diethyl ether was added, and the mixture shaken. Then 0.8 volumes of 0.33 M NaCl was added, the mixture shaken again, and the layers allowed to separate. The upper, pigmented ether phase was decanted, dried by the addition of anhydrous sodium sulphate, and the ether phase decanted. The ether was then evaporated under vacuum.

Carotenoids were extracted from washed chloroplasts in the dark by resuspending the pellet in 50 mL of 96% ethanol buffered with sodium carbonate and incubating on ice. Unextracted material was removed by centrifugation at 12,000 g for 15 min at 4°C. To achieve saponification, 80 % KOH (w/v) was added and the mixture incubated at 4°C overnight in the dark. The solution was placed in a separating funnel and mixed gently with one volume of diethyl ether, and then with 0.8 volumes of 0.33 M NaCl. After distinct layers had formed the lower green phase was separated from the orange upper ether phase, to which three volumes of water were then added with gentle mixing. After phase separation the lower phase was removed, and the top phase dried with anhydrous sodium sulphate. The ether was then evaporated under vacuum.

Extracted pigments were re-dissolved in 100 % acetone and their concentration was checked by absorbance spectroscopy in 80 % buffered acetone (using 20 mM Tris (pH 8.0)). Concentrations of Chl *a*, Chl *b* and total carotenoid were calculated based on published equations[72,181]. The Chl *a*:Chl *b* ratio was near 3.0 as expected. Total pigments and carotenoid extracts were aliquoted into 1 mg and 160 μg amounts and freeze-dried. The dried pigments were stored at -80°C before use.

2.6.3. LHCII reconstitution and purification

The protocol used for LHCII reconstitution was based on that described in JoVE[72], available online. The main variation was that ß-mercaptoethanol was replaced by DTT.

LHCI complexes were reconstituted by mixing apoprotein from IBs with an appropriate ratio of pigments that were thoroughly pre-dissolved in ethanol in the presence of lithium dodecyl sulphate (LDS), and then LDS was replaced by OG using KCl to precipitate the LDS.

A 800 µg sample of IBs were resuspended in 400 µL TE and mixed with 400 µL of 2x Reconstitution Buffer (200 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)/5% (w/v) sucrose/4% (w/v) LDS/2 mM benzamidine/10 mM aminocaproic acid) by brief vortexing. DTT was added to a final concentration of 10 mM and the protein heated for 1 min at 98°C before vortexing briefly and incubating at RT for 3 min.

A 500 µg aliquot of total dried pigments and 80 µg of dried carotenoid were vortexed for 1-2 min in 30 µL of 100% ethanol. The pigments were centrifuged at 16,000 g at 4 °C for 30 s to check that suspension was complete. The pigment was immediately added slowly to the cooled protein while vortexing, and vortexing was continued for 10 s before placing on ice. OG was then added to a final concentration of 2% and the tube incubated on ice for 10 min. KCl was then added to give a final concentration of 200 mM, the solution vortexed briefly and placed on ice for 20 min. The solution was then centrifuged for 10 min at 16,000 g at 4 °C and the supernatant removed without disturbing the precipitated LDS pellet. The solution was kept on ice in the dark.

Reconstituted pigment-protein was purified by nickel-affinity column chromatography and sucrose gradient ultracentrifugation as described elsewhere[72]. Further purification of the protein was achieved by gel-filtration chromatography in Tris-DDM. Fractions were kept and pooled if they had a low A_{470} to A_{674} ratio and identical emission profiles under 440 nm, 475 nm and 500 nm excitation, specific for Chl *a*, Chl *b* and xanthophyll carotenoids, respectively. Purified proteins were concentrated and stored at -80°C before use.

The quality of reconstitution was checked by absorbance and emission spectroscopy of the final product and absorbance spectroscopy of an 80% acetone extract. Final Chl *a*:Chl *b* ratios and total Chl:protein ratios were calculated for all reconstituted complexes. An extinction coefficient at the Chl *a* Q_y band equal to 546,000 M^{-1} cm^{-1} was used to estimate the concentrations of LHCII.

2.6.4. Spinach LHCII extraction and purification

Thylakoid membranes were purified from fresh spinach leaves following a previously published protocol[72]. The concentration of membrane in terms of total Chl was then calculated and membranes at 0.5 mg mL^{-1} Chl were solubilized with 1.2 % DDM. Linear sucrose gradients were prepared by freeze-thaw using a 21 % (w/v) sucrose solution in 10 mM HEPES (pH 7.6)/0.03 % DDM, loaded with solubilised membrane and centrifuged for 18 hrs at 38,000 rpm in a Sorvall TH-641 swing-out rotor. The composition of each fraction taken from the resulting sucrose gradients was examined by SDS-PAGE, the results compared to published data and the top fraction containing mostly LHCII was collected [75,182]. This was then subjected to gel-filtration chromatography to separate LHCII from trace denatured proteins and LHCII-PSII complexes. Eluted fractions with the lowest A_{470}:A_{674} nm ratio and emission profiles that were invariant with respect to excitation wavelength were pooled.

2.6.5. LHCI modification, refolding and purification.

LHCI heterodimeric complexes were assembled from modified versions of *A. thaliana* proteins Lhca1 (UniProtKB entry Q01667) and Lhca4 (UniProtKB entry P27521). Expression plasmids were pET-28a containing synthetic genes sourced from Eurofins. Following apoprotein expression in *E. coli* Rosetta™ 2 (**Table 2.1**), LHCI heterodimers were assembled by refolding with purified pigments in DDM using the same protocol as for LHCII. A 20% excess (by mass) of either L1 or Td-L1 was mixed with Td-L4 to ensure a minimal level of free Td-L4 monomer after refolding. The apoprotein:total pigment ratio was kept the same as for LHCII refolding. Nickel affinity chromatography was used to separate the His-tagged LHCI dimer from residual Lhca1 monomer (which was not His-tagged). Each LHCI was then further purified by gel filtration chromatography and stored at -80°C before use as a concentrated solution in Tris/DDM. An extinction coefficient for the Chl *a* Q_y band equal to 1,092,000 M^{-1} cm^{-1} was used to evaluate LHCI concentration since its Chl *a* content is approximately twice that of a refolded LHCII monomer [73].

2.7. XylE expression and purification

The gene for the *E. coli* proton-coupled xylose transporter XylE modified with a sequence encoding a His-tag was cloned into a pET28 vector and expressed in *E. coli* strain C43(DE3). Expression was induced in 1 L cultures at a cell density of OD_{600} = 0.8 by adding 1 mM IPTG for 3 h during growth in an orbital incubator at 250 rpm and 37°C. Harvested cell pellets were resuspended in 25 mM Tris (pH 8)/150 mM NaCl and a cOmplete™ EDTA-free protease inhibitor tablet and a few crystals of lyophilised DNase (Sigma Aldrich®) were added. Cells

were lysed by two passes through a Constant Systems cell disrupter at 25,000 psi. Cell debris was removed by centrifugation 15,000 rpm for 15 min at 4°C in an SS-34 rotor and the membrane fraction isolated by centrifugation of the supernatant at 38,000 rpm for 1 h at 4°C in a Ti45 rotor. The resulting membrane pellet was resuspended in 25 mM Tris (pH 8)/150 mM NaCl and DDM was added to a final concentration of 1.5 % (w/v). After stirring for 1 h at 4°C membrane debris was removed by centrifugation at 38,000 rpm for 1 h at 4°C in a Ti45 rotor. The supernatant was diluted in 25 mM Tris-HCl (pH 8)/150 mM NaCl to achieve a final concentration of 1 % DDM and passed through a 5 mL Ni-NTA column (GE Healthcare). Bound His-tagged XylE ($XylE_H$) was eluted using 25 mM Tris (pH 8)/150 mM NaCl/300 mM imidazole/0.05% DDM. The eluate was concentrated and the $XylE_H$ was further purified by gel-filtration chromatography using a Superdex 200 column equilibrated in 25 mM Tris (pH 8)/150 mM NaCl/0.05% DDM. The $XylE_H$ peak was collected, concentrated, aliquoted, snap-frozen in liquid nitrogen and stored at -80°C. The concentration of $XylE_H$ was evaluated using a *DC*™ protein assay (Bio-Rad).

2.8. eGFP/MBP soluble protein expression and purification

Expression of engineered versions of eGFP and maltose binding protein (MBP) in pET28a expression vectors was achieved using *E. coli* strain Rosetta™ 2 grown at 37°C in LB medium until OD_{600} reached 0.6 ~ 0.8. The induction time, temperature and IPTG concentration were optimized to maximize the yield, but were typically between 0.4 mM to 1 mM IPTG at a temperature of 21 °C to 37 °C for 4 ~ 6 h depending on the difficulty of expression. Cells were lysed in the same way as for RC purification but under 25,000 psi pressure. The lysed cells were immediately subjected to ultracentrifugation to remove any debris and membranes and supernatant was then loaded onto a 5 mL Ni-NTA column equilibrated in 20 mM Tris (pH 8.0)/200 mM NaCl/20 mM imidazole. The bound protein was eluted using 20 mM Tris (pH 8.0)/200 mM NaCl/500 mM imidazole and further cleaned and buffer exchanged into 20 mM Tris (pH 8.0) using a Superdex 200 gel-filtration column.

2.9. Peptide synthesis and characterization

2.9.1. Peptide synthesis

Peptides were synthesised by the solid-phase assist method. Peptides JL1 and JL2 were synthesized on a CEM Liberty Blue™ Automated Microwave-assisted Peptide Synthesizer using Fmoc-amino acids in DIC/Cl-HOBt/DMF solution (HOBt - 4-formyl-3-methoxyphenoxy)butanoic acid, 1-hydroxybenzotriazole; DIC - *N,N'*-diisopropylcarbodiimide; DMF - dimethylformamide). Deprotection was achieved by 5% piperazine/0.1 M formic acid in

DMF. After synthesis of peptides on support amide-resins, the resins were split into halves. Peptides were capped and cleaved from the resin and freeze dried.

2.9.2. Peptide purification and characterization

Reverse-phase high pressure liquid chromatography (HPLC) was used for peptide purification using a 20%-80% acetonitrile gradient. The major peaks at 280 nm absorption were collected in ~0.5 mL volumes and the composition of fractions was evaluated using an Applied Biosystems 4700 Proteomics Analyser (matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry). Samples were mixed with 20 mg mL^{-1} α-cyano-4-hydroxycinnamic acid in equal volumes and dried onto a MALDI plate. The correct molecular weights for peptides JL1 and JL2 should be 3744 g M^{-1} and 3436 g M^{-1}. Fractions with mostly single correct peaks were pooled and freeze dried. The purified peptides were re-dissolved in water and evaluated again by reverse-phase HPLC. Concentrations were determined by absorption at 274 nm for JL1 and 214 nm for JL2. The extinction coefficient used for JL1 (274 nm) was 1400 $M^{-1}cm^{-1}$ in water and 39708 $M^{-1}cm^{-1}$ for JL2 (214nm) in 20 % acetonitrile/0.1% formic acid.

2.10. Polyacrylamide gel electrophoresis

2.10.1. SDS-PAGE

SDS-PAGE analysis was carried out using 4 – 20% Mini-PROTEAN® TGX™ Precast Protein Gels (Bio-Rad). Typically about 20 pmol RC was used in each lane, and gels were subjected to 200 V for about 45 mins. Two methods were used to stain protein, quick Coomassie stain (Generon™) and SYPRO™ Ruby stain (Thermo Fisher). Quick Coomassie staining was done by incubating the gel with quick staining reagent overnight. A slow SYPRO Ruby stain protocol was used for better quantification of protein composition. Washed gels were visualized using a Gel Doc™ EZ System (Bio-Rad).

2.10.2. Blue Native (BN) PAGE

The protocol used was adapted from a published procedure[183]**.** The cathode buffer contained 0.02% (w/v) Coomassie blue dye and there was no dye in loading buffer. Precast NativePAGE™ 4-16% Bis-Tris gels were run in a Bris-Tris buffer system (Thermo Fisher). The whole gel cassette was placed in ice bath running at 150 V for 1 h and then raised to 250 V for about 2 h. After the dye front started to leak into the anode chamber, the gel was removed

and destained in 10% (w/w) methanol and 7% (w/w) acetic acid solution until a good background was seen.

2.10.3. Western Blot

Partial protein was transferred to nitrocellulose membrane (GE Healthcare) using a TE 77 PWR Semi Dry Transfer Unit (GE Healthcare) run at 45 mA gel^{-1} and 30 min using a NOVABlot Kit. The membrane was then blocked overnight with 5 % (w/v) milk powder in phosphate buffered saline (PBS)/0.01 % Tween-20 (PBS/T buffer) and then incubated with horseradish peroxidase (HRP) conjugated antibodies in the same buffer for 1 h. The membrane was developed using 1x LumniGLO® reagent (Cell Signalling Technology) after rinsing the membrane three times with PBS/T buffer. Finally, the result was recorded on an ODYSSEY imaging system (LI-COR Biosciences). Re-probing of the membrane was accomplished by stripping and repeat of process from 5 % Milk PBS/T incubation. Stripping of membrane was achieved by incubating in mild stripping buffer (200 mM glycine (pH 2.2)/0.1% SDS (w/v)/1% Tween 20 (v/v)) for 5 mins twice and two times in TBST buffer (50 mM Tris (pH 7.6)/150 mM NaCl, 0.1% Tween 20 (v/v)) and finally transferred in PBS/T.

2.11. Quantum yield determination

2.11.1. QD and XylE/QD conjugate quantum yield

Water-soluble CdTe QDs coated with 3-mercaptopropionic acid (MPA) with an emission maximum at 750 ± 5 nm were purchased from PlasmaChem GmbH. The supplied average molecular weight of these QDs (550 KDa) was used to calculate their molar concentration and molar extinction coefficients.

The quantum yield of the water-soluble QDs (Φ_{QD}) in Tris/DDM was determined by comparison of their emission spectra to that of the dye LDS-751 (laser dye styryl–751, Sigma) dissolved in methanol.[184] Potential effects of protein attachment were accounted for by incubating QDs with the His-tagged, DDM-solubilised photochemically-inactive membrane protein $XylE_H$ (Section 2.7) at the same ratios as used for RC/QD measurements. To avoid self-shading the absorbance of QD/XylE conjugates and LDS-751 was limited to 0.025 at the 550 nm excitation wavelength. Emission from LDS-751 and QDs or QD/XylE conjugates after 550 nm excitation (average of 10 measurements) was corrected with spectral response and used to calculate the relative integral photon fluxes of QDs and LDS-751.[185] The value for Φ_{QD} was estimated with reference to $\Phi_{LDS\text{-}751}$ = 0.014 with the value of refractive index of water

(n_{water} = 1.333) and methanol ($n_{ethanol}$ = 1.328)[WPC18]. The estimated Φ_{QD} for XylE/QD conjugates was constant over the range of protein to QD ratios examined and the average value of 0.197 was used for calculating FRET distance.

2.11.2. LHCII quantum yield

The quantum yield of dLHCII (Φ_{LHCII}) in Tris/DDM was determined by comparison of its emission properties to those of the dye 4-dicyanomethylene-2-methyl-6-p-dimethylaminostyryl-4H-pyran (DCM - Sigma) dissolved in methanol. To avoid self-shading the absorbance of dLHCII conjugates and DCM was adjusted to below 0.1 across the relevant spectral region. Emission of DCM and dLHCII (average of 10 measurements) was corrected with spectral response and used to calculate the relative integral photon fluxes of dLHCII and DCM.[185] The value for Φ_{LHCII} was estimated with reference to Φ_{DCM} = 0.435 [186] and the value of refractive index of water (n_{water} = 1.333) and methanol ($n_{methanol}$ = 1.328). Estimated Φ_{LHCII} was approximately 0.135(SD < 0.01) which was consistent with literature values.[68,131]

2.11.3. Quantum yields of other LHCII or LHCI

Quantum yields of other engineered LHCII and LHCI pigment-proteins were determined by comparison to the quantum yield of dLHCII that was established relative to DCM. All emission profiles were corrected by absorbance over the excitation range and total photon counts were used for the comparison.

2.12. Transmission Electron Microscopy

2.12.1. Grid preparation

For RC-LHCII chimeras, negative stain transmission electron microscopy (TEM) was carried out on an equimolar mixture of 500 nM WT RCs and dLHCII or 500 nM LHCII#RC heterodimers. For conjugates, negative stain TEM was carried out on conjugates formed from a 10:1 RC:QD mix or 5:5:1 LHCII:RC:QD mix at a QD concentration of 100 nM. In all cases a drop of sample was incubated on a glow discharge treated carbon coated grid for 30 s. Filter paper was used to remove excess, the grid was floated on top of a 3 % uranyl acetate (Sigma) droplet, and filter paper was used for a second time to remove excess liquid. After one repeat of this procedure the grid was placed for one minute on a third droplet of uranyl acetate. The grid was completely dried in air before imaging with a Tecnai 12 120kV BioTwin Spirit TEM in the Wolfson Bioimaging Facility, UoB.

2.12.2. Post imaging processing and analysis

Recorded images were analysed in Matlab (MathWorks). To increase the rate of detection of true conjugates and reduce false selection, images were de-noised with a Wiener filter and the contrast enhanced using the Matlab CLAHE (contrast-limited adaptive histogram equalization) algorithm. Conjugate selection was carried out on a binary gradient image by detecting continuous edges outlining the shape of potential conjugates. Most identified particles were approximately spherical in shape and their diameters were in the physically permitted range. A final manual check was carried out to exclude any objects with a false outline, and particle diameters were fitted to a lognormal distribution with an upper cut-off of 33 nm.

2.13. Dynamic light scattering

Particle sizes were measured by dynamic light scattering (DLS) on a Zetasizer Nano ZSP (Malvern Instruments Ltd) kindly made available by Dr. Paul Curnow. Samples were pre-equilibrated at 25^{o}C for 5 mins and data were collected at 173^{o} forward scattering with 5 repeats. Concentration of sample was set to 100 nM to minimize the emission and absorption affect light scatter signals. Average particle sizes were computed from the intensity distribution peaks.

2.14. Sucrose density gradient ultracentrifugation

2.14.1. Sucrose cushion for membrane isolation

Isolation of photosynthetic membranes was carried out using a sucrose cushion method to minimise aggregation. Firstly, cell lysates were prepared as described in section 2.5.1. After a clearing centrifugation spin the lysate was loaded onto a 60% sucrose layer, which provided a cushion about 2 – 3 cm from the bottom of the ultracentrifuge tube. After centrifugation at 38,000 rpm for 2 hrs in 4^{o}C with a Ti-45 rotor the membrane formed a tight band at the 0%/60% sucrose interface. This band was collected and dialysed against 20 mM Tris (pH 8.0) to remove sucrose. Final isolated membranes were stored at -20^{o}C after aliquoting and snap-freezing.

2.14.2. Continuous sucrose density gradients for separation of RCs, LHCs and chimeras

Linear sucrose density gradients were prepared by placing 10 mL aliquots of 21% (w/v) sucrose in TrisDDM in 12 mL ultraclear ultracentrifuge tubes and freezing and thawing these once. Each gradient was then loaded with 400 μL of sample (typically at a photoprotein concentration of 2.5 μM) and then capped with 1 mL buffer. Gradients were ultracentrifuged in a TH-641 swing-out rotor at 180,000 g for 18 h at 4°C. To assess the composition of protein along the gradient, either bands were withdrawn from the top of the tube in 1 mL volumes, or they were collected from the bottom of the tube by puncturing a hole at the bottom and collecting 1 mL fractions dropwise.

2.14.3. Two-step sucrose density gradients for separation of proteins and conjugates

RC/QD, LHCII/QD or LHCII/RC/QD conjugates were separated from free proteins on two step discontinuous sucrose density gradients formed from equal volumes of 25 % and 60 % (w/v) sucrose in Tris/DDM in 12 mL ultraclear ultracentrifuge tubes. A standard loading of 400 µL of sample was used with a varying concentration of QDs and 2 mM tris(2-carboxyethyl)phosphine (TCEP - Thermo®) was present when LHCII was used. Samples were then overlaid with 1 mL of Tris/DDM to form a second step. Gradients were centrifuged at 38,000 rpm at 19°C for 4 hours using a TH-641 swing-out rotor and deconstructed in ten 1 mL fractions and one 1.4 mL fraction by the dropwise method (see above). Fractions 9-11 from the top of the gradient corresponded to free proteins. Fractions 1-8 from the bottom and middle of the gradient corresponded to protein/QD conjugates and a small population of proteins released from QDs during the fractionation run. The structural integrity of proteins in fractions 9-11 were assessed by absorption spectroscopy or SDS-PAGE and Western blot (Section 2.10).

2.15. Fluorescence titrations

Fluorescence titrations were carried out on a Cary Eclipse Fluorescence Spectrophotometer (Agilent™) with a microplate attachment, using black 96-well microplates (Greiner Bio-One™) and a standard sample volume of 100 µL. The spectrophotometer light source was a pulsed xenon lamp operating at 80 Hz and with a pulse full width at half maximum (FWHM) of 2 µs. Titrations of LHCII and QD emission were conducted in dilute solutions in which the QD concentration was fixed at 50 nM with varying concentrations of protein. The absorbance spectra of all samples were also recorded. Emission spectra were baseline corrected, then individual LHCII and QD emission spectra were deconvoluted using reference spectra taken within the same experiment to distinguish QD emission peaking at 750 nm and LHCII emission peaking around 682 nm. Some measurements involving LHCs

were carried out in nitrogen-gassed Tris/DDM supplemented with 200 µg/mL glucose oxidase, 7.5 mg/mL glucose, and 35 µg/mL catalase to scavenge oxygen.[187]

For titrations with fluorescent VL157R RCs, baseline-corrected spectra were deconvoluted to distinguish emission from the QDs at 750 nm and the RC at 800 nm. Peak values were then used to represent the fluorescence from each component. Absorbance spectra for all samples were recorded on a Cary UV-Vis spectrophotometer and were used to correct the emission signal. For comparison of samples with the same emission profile no spectral response correction was made, whereas when comparing different emitting materials, the spectral response was corrected using a correction factor generated using reference dyes.

2.16. Charge separation in RCs

Photo-oxidation of P870 was measured using an optical fibre attachment for a Cary 60 UV-Vis spectrophotometer and a four-way cuvette holder (Ocean Optics, Inc.). For excitation, light from a HL-2000 tungsten halogen source (Ocean Optics, Inc.) was passed through an optical fibre and a 430 ± 12.5 nm or 650 ± 12.5 nm bandpass filter (Edmund Optics Ltd., catalogue numbers 86653 and 86657 respectively). Incident light intensity was approximately 0.3 mW, which yielded $P870^+$ in 5 – 15 % of the RC population. Light-on/off was controlled using the electronic shutter on the light source triggered by a TGP110 pulse generator (Aim-TTi Ltd, United Kingdom). After incubation with a 10-fold excess of ubiquinone-0 (UQ_0) in the dark for 2 min, samples at a total protein concentration around 1 µM were housed in a 3 mm path length, four-sided micro cuvette (110-15-QS, Hellma® Analytics). P870 oxidation was induced by a seven second period of continuous illumination. Each measurement was repeated five times and averaged traces were fitted to a model assuming a simple interconversion between the ground (P) and photo-oxidised state (P^+) of P870:

$$P \underset{k_r}{\overset{k_f}{\leftrightarrow}} P^+ \qquad (\mathbf{2.1})$$

where k_f is positively correlated with the energy utilised by the RC.

The recovery of RC ground state absorbance after white light illumination for 0.5 s was also measured with the same setup minus the bandpass filter. Recovery rates were determined from traces that were the average of five repeats.

2.17. Molecular modelling

2.17.1. Design of SpyCoil using ISAMBARD and Modeller

A model coiled-coil dimer was generated in ISAMBARD[160] using the P3/P4 sequences designed by Bromley *et al*.[148] and using SCWRL4.[188] Using this package, the joint between SpyTag and helix P3 was also finished with one or two glycines. The SpyTag/KTag structure was derived from the SpyTag/SpyCatcher crystal structure reported by Li *et al.*[156] A SpyLigase was also generated with the missing residues using the Modeller[189] loop refinement function in the presence of SpyTag and KTag to avoid backbone clash. The partially complete SpyCoil assembly was freely rotated at the glycine flexible linker and the free energy of system at various states were estimated using Bristol University Docking Engine (BUDE) to validate the design. Then the missing loop attaching KTag and P4 was generated in Modeller with various numbers of amino acids. The shortest linker that did not pose structural disorder to KTag and P4 α-helix was accepted as the final construct.

2.17.2. Model construction with SCWRL and Modeller

Construction of model proteins with various protein motifs and domains were generated in Modeller with the loop refine model with at least five repeats. Energy scores were calculated in discrete optimized protein energy (DOPE) scores to validate the model. All structural models presented in this thesis were derivatives of published X-ray crystal structures with missing residues and any additional sequences added to the host proteins.

2.17.3. Molecular dynamics with GROMACS

Molecular dynamics (MD) of designed SpyCoil proteins were accomplished with GROMACS.[190] After OPLS-AA/L all atom force field was used as the potential function of selected proteins, protein was included in a cubic box with at least 2.5 nm to the box edge and spc216.gro was used to configure solvent. The charge of system was balanced with Na^+ and Cl^-. After minimising the system with a short robust MD, the system was equilibrated to a temperature equal to 300 K and pressure equal to 1.0 bar. Each equilibrated system was run for 50 ns on the supercomputer BlueGem (8 core/16 nodes). The result protein trajectories were extracted for analysis.

2.17.4. Analysis of surface coverage effects by Monte Carlo 2D packing of hard particles

Based on the new polygon functions in MATLAB 2018, random packing of either covalently linked RC-LHCII or mixtures of RC and LHCII was performed with the 2D projection of protein perpendicular to the membrane plane plus 1.5 nm DDM micelle[191]. First, small objects (0.2 of original) were randomly deposited to the field with a periodic boundary and then check for overlaps. After resolving the overlaps, the particle size was increased by 0.08%, then a new round of fitting started until the particle grew to its original size. For separated mixture of proteins, two parameters – mass centre coordinate (*C(x,y)*) and orientation of protein regarding reference axis (θ) were randomly assigned and any overlap was solved with small perturbations on *C(x,y)* and θ. If overlap could not be resolved by small step perturbation, then a new set of parameters was given to the enquired object, which was subjected to the next round of perturbation. For packing of covalently attached RC#LHCII, two more parameters were introduced described the distance (*d*) and the relative angle of the two proteins (φ). *d* was varied in the permitted range determined by flexible linker on two proteins and the SpyCatcher/SpyTag complex (0 to 2.9 nm). Appropriate values of *d* and φ were accepted without any overlap between RC and LHCII discs, then the conjugate was looped to eliminate overlap with other conjugates as described above until the size relaxed to the original size. Then the initial field was relaxed with small steps (0.02 of RC maximal diameter) until the self-dynamic overlap function converged to zero.[192]

2.18. Photochronoamperometry and EQE action spectra.

Nanostructured silver electrodes of 2 mm depth were prepared as described previously.[140] Pigment-proteins at concentrations between 20 µM and 100 µM were drop-casted onto prepared electrodes in the dark at 4 °C for one hour and unbound protein was removed by repeated mechanically-controlled dipping in 20 mM Tris (pH 8) at 4 °C. Coated electrodes were immersed in 20 mM Tris (pH 8)/50 µM KCl/20 µM cyt *c*/1.5 mM ubiquinone-0 (Q_0) in a room temperature electrochemical cell fitted with an Ag/AgCl/3M KCl reference electrode and a platinum counter electrode. Photocurrents were measured at a bias potential of -50 mV vs Ag/AgCl, controlled by a PGSTAT128N potentiostat (Metrohm Autolab). Illumination was supplied by 870 nm or 680 nm light emitting diode (LED - Roithner Lasertechnik) with irradiances of 32 or 6.7 mW cm^{-2}, respectively, at the electrode surface (~50 nm FWHM for both). External quantum efficiency (EQE) action spectra were recorded using a tungsten-halogen or xenon light source passed through a monochromator, as described previously.[140]

Chapter 3

Mechanisms of Self-assembly and Energy Harvesting in Conjugates of Cadmium Telluride Quantum Dots and the *Rhodobacter sphaeroides* Reaction Centre

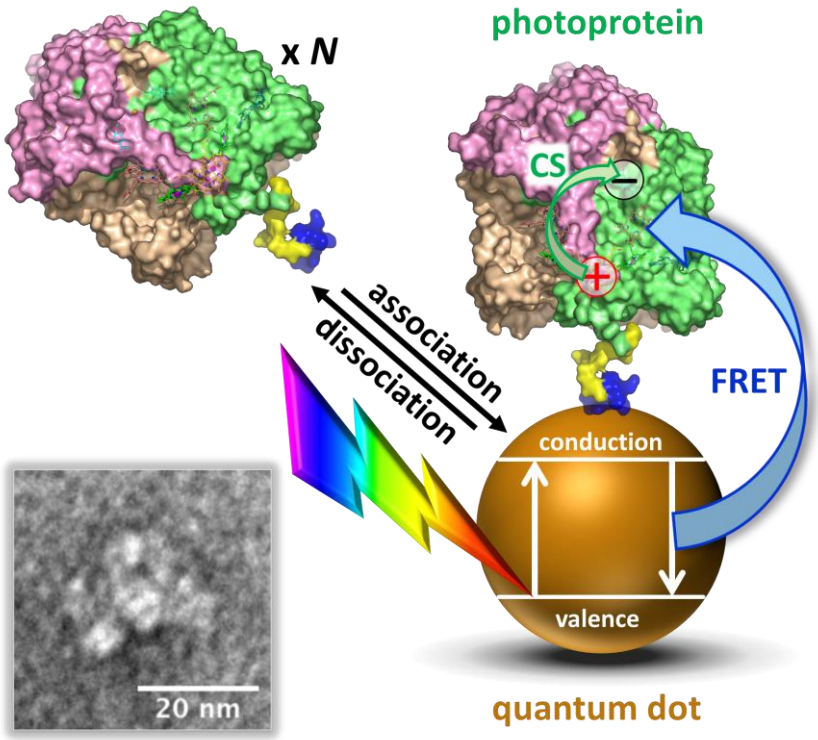
A manuscript describing the results presented in this chapter has been published and was written by me.

Liu, J., Mantell, J., Di Bartolo, N. and Jones M.R. (2019) Mechanisms of self-assembly and energy harvesting in tuneable conjugates of quantum dots and engineered photovoltaic proteins. *Small* 15, 1804267
https://doi.org/10.1002/smll.201804267

The experiments and analysis described in this chapter were performed by me. Dr. J. Mantell assisted with the collection of transmission electron microscopy images. Dr. N. Di Bartolo supplied His-tagged XylE.

Abstract

Photoreaction centres facilitate the solar energy conversion at the heart of photosynthesis and there is increasing interest in their incorporation into biohybrid devices for solar energy conversion, sensing and other applications. In this chapter the self-assembly of conjugates between engineered *Rba. sphaeroides* RCs and QDs that act as a synthetic light harvesting system is described. The interface between protein and QD is provided by a poly-histidine tag that confers a tight and specific binding and defines the geometry of the interaction. Protein engineering that changes the pigment composition of the RC is used to identify Förster resonance energy transfer (FRET) as the mechanism through which QDs can drive RC photochemistry with a high energy transfer efficiency. A thermodynamic explanation of RC/QD conjugation based on a multiple/independent binding model is provided. The results also demonstrate that the presence of multiple binding sites affects energy coupling not only between RCs and QDs but also among the bound RCs themselves, effects which likely stem from restricted RC dynamics at the QD surface in denser conjugates. These findings are readily transferrable to many other conjugate systems between proteins or combinations of proteins and other nano-materials.



3.1. Introduction

A feature of natural photosystems is selective harvesting of certain regions of the solar spectrum, the most obvious example being the predominant green colour of plant photosynthetic tissues that arises from relatively strong absorbance of red and blue light by chlorophyll and carotenoid pigments. As *Rba. sphaeroides* synthesises BChl *a* as its primary photosynthetic pigment its RC (**Fig. 3.1A,B**) exhibits strong absorbance in the near-infrared between 700 and 950 nm, and in the near-UV below 420 nm, but its absorbance across the visible region is relatively weak (**Fig. 3.1C**). A limitation in the use of this protein in device technologies is therefore sub-optimal harvesting of light energy across much of the region where the solar radiation at the earth's surface is maximal,[125] and this limitation is manifest in action spectra of photocurrent density in photoelectrochemical cells based on *Rba. sphaeroides* pigment-proteins[134,136,139,140,193,194].

In this chapter I investigated directed self-assembly of well-defined conjugates between genetically-engineered RCs and water-soluble CdTe QDs. The QDs have broad absorbance across the visible spectrum and an emission band centred at 750 nm that overlaps with RC absorbance bands centred at 760 nm and 800 nm (**Fig. 3.1C**). These QDs were therefore capable of acting as a synthetic light harvesting system for energy transfer[195,196] and charge separation[38,51,52,57] in the *Rba. sphaeroides* RC (**Fig. 3.1B**).

The aim of the work was to find out if energy transfer between photo-excited QDs and bound RCs can be explained by FRET[150,197–199] and to discover whether a thermodynamic description of self-directed binding between RCs and QDs based on a multiple/independent binding model[200] could be developed. The latter was important in order to address the challenge of understanding the properties of the tri-component conjugates described in Chapter 4. QDs are uniform structures and any portion of the surface can act as a binding site. They could therefore accommodate multiple RCs and hence energy capture could be tuned up to a 92 % efficiency in the conjugates by varying the ratio of RCs to QDs. In conjugates formed between a cofactor-modified RC and QDs, increasing the RC density also caused the appearance of a new energy quenching pathway, suggesting that controlling protein density can also be used as a way of switching functionality.

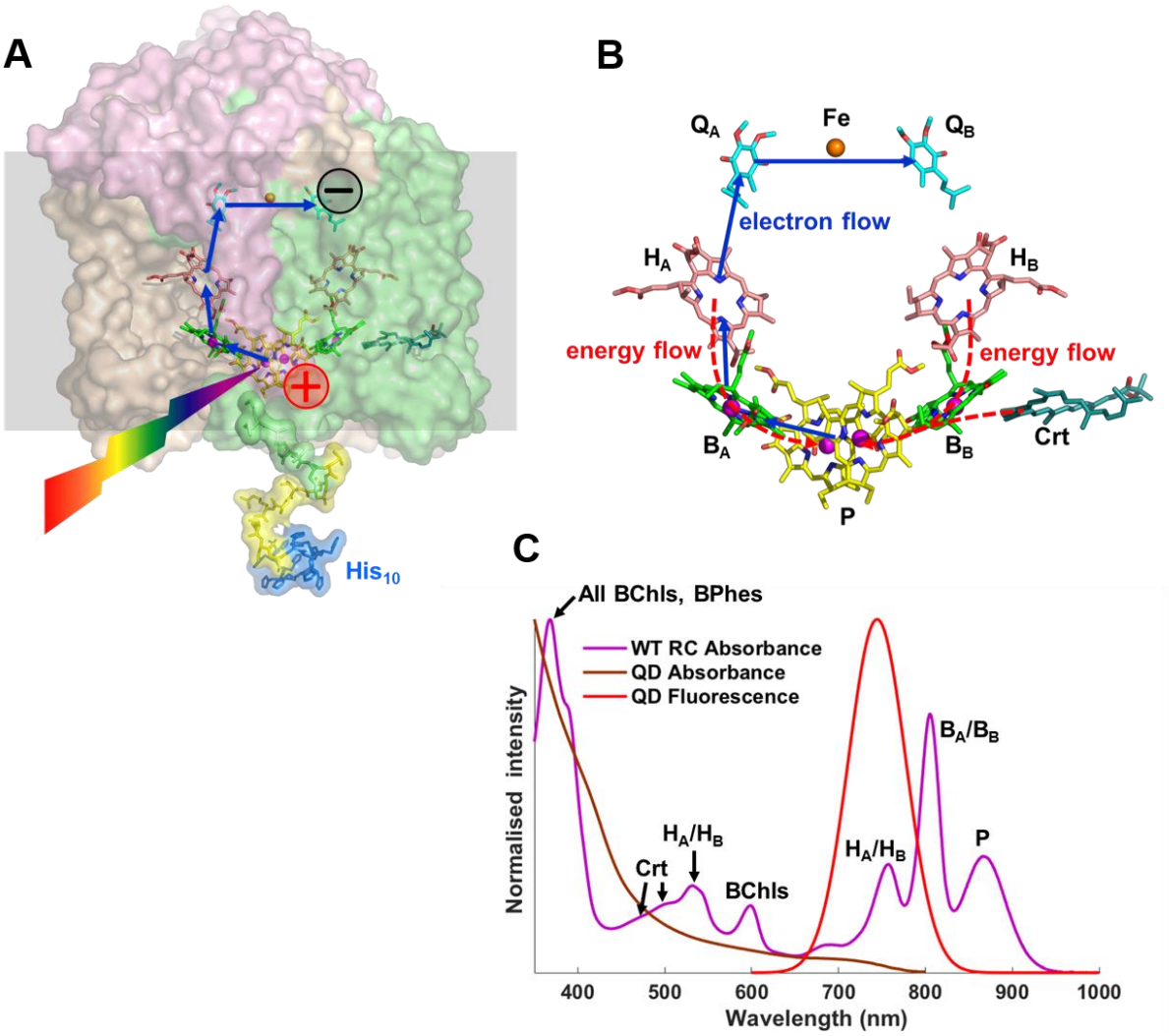


Figure 3.1. Structure of the RC and optical properties of components. (A) The *Rba. sphaeroides* RC comprises three polypeptides (transparent pink, green and beige surfaces) that scaffold ten cofactors embedded in the interior of the photosynthetic membrane (grey box). Photo-excitation separates charge between BChl and ubiquinone cofactors on opposite sides of the membrane (blue arrows). An engineered extra-membrane deca-histidine tag (blue surface) is connected to the protein by a linker (yellow surface). (B) Excited states formed on the RC BPhe (H_A/H_B – pink carbons)) or monomeric BChl (B_A/B_B – green carbons) cofactors by direct photon absorption or energy transfer from an antenna migrate to the P BChl pair (yellow carbons) on a sub-picosecond time scale (red dashed arrows). The resulting P* excited state triggers charge separation (blue arrows). (C) CdTe QDs absorb across the visible region and their emission overlaps with the lowest energy RC absorbance bands in the near-IR.

3.2. Results

3.2.1. Methodologies

Wild-type (WT) and cofactor-altered RCs, either otherwise unmodified or modified with a deca-histidine tag on the C-terminus of the M-polypeptide, were expressed in *Rba. sphaeroides* strain DD13 lacking light harvesting complexes (see **Table 2.2** in Chapter 2). The procedures used for the growth of *Rba. sphaeroides* are described in Section 2.2.1. RCs without His-tags were purified by ion exchange chromatography (Section 2.5.3) while His-tagged RCs were purified by nickel affinity chromatography (Section 2.5.2). The His-tag was placed at the C-terminus of PufM, as described previously[113] and RCs with this modification were denoted with a subscript "H". RC modifications were carried out as described in Section 2.4.

Water-soluble Cd/Te QDs coated with 3-mercaptopropionic acid with an emission maximum at 750 ± 5 nm were purchased from PlasmaChem GmbH. The supplied average molecular weight of these QDs (550 KDa) was used to calculate their molar concentration.

3.2.2. Mechanism of binding of RCs to QDs.

Sucrose gradient ultracentrifugation using two-step 0%/25%/60% gradients (Section 2.15.3) was used to investigate the mechanism of binding of WT RCs to 750 nm-emitting water-soluble CdTe QDs (**Fig. 3.2A**). In single component samples, QDs migrated to the 25/60 % interface in the lower part of a two-step sucrose gradient, whereas RCs remained at the upper 0/25 % interface. When RCs and QDs were mixed in a 10:1 ratio, ~95 % of RCs were pulled down to the lower interface only when modified with a His_{10}-tag (**Fig. 3.2A**). Henceforth in this chapter WT RCs modified with a His-tag are denoted WT_H.

The conclusion that binding of RCs to QDs was mediated by the protein His-tag was supported by the finding that WT RCs quenched QD emission only when His-tagged (**Fig. 3.2B** and **Fig. 3.3A**). Details on emission spectroscopy are given in Section 2.15. Measurements taken over extended time periods showed that this quenching was due to a stable association with His-tagged WT_H RCs at all RC:QD ratios tested (**Fig. 3.3B**). Quenching could be reversed by addition of imidazole or histidine, but not by NaCl (not shown). Cleavage of the His-tag using thrombin also produced recovery of QD emission that was dependent on the extent of cleavage (**Fig. 3.3C**).

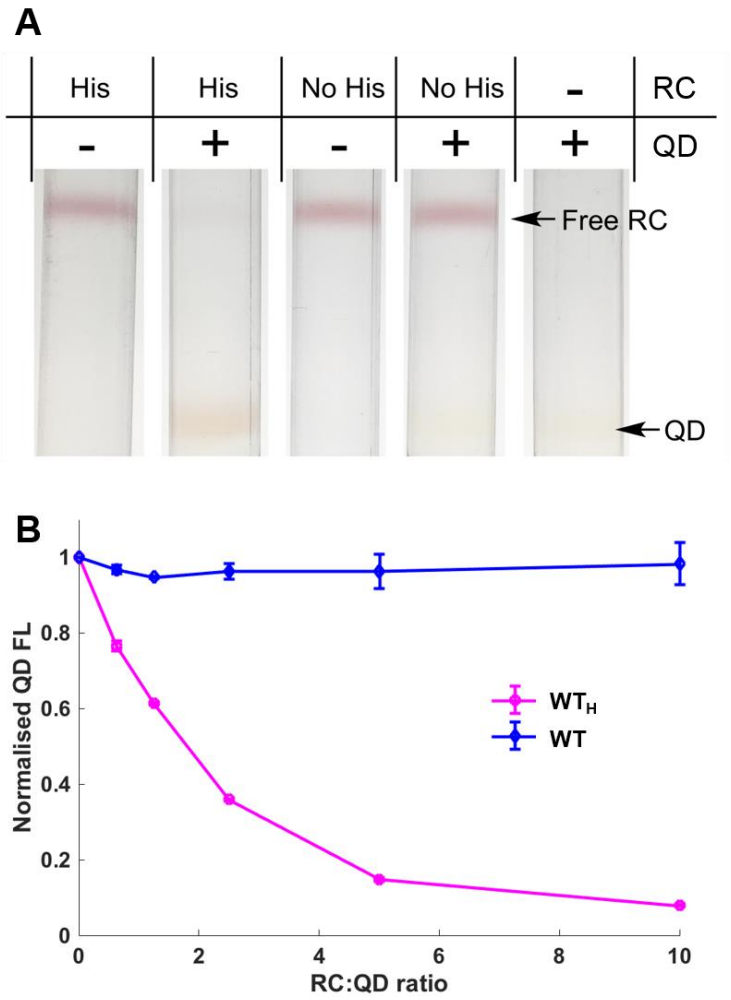


Figure 3.2. Binding of RCs to QDs and quenching of QD emission. (A) Two-step sucrose gradients that separate unbound RCs (top red band) from QDs or RC/QD conjugates (bottom brown band). (B) Intensities of QD emission at different RC:QD ratios relative to a QD-only control (excitation at 430 nm), showing that only His-tagged WT RCs quenched QD emission.

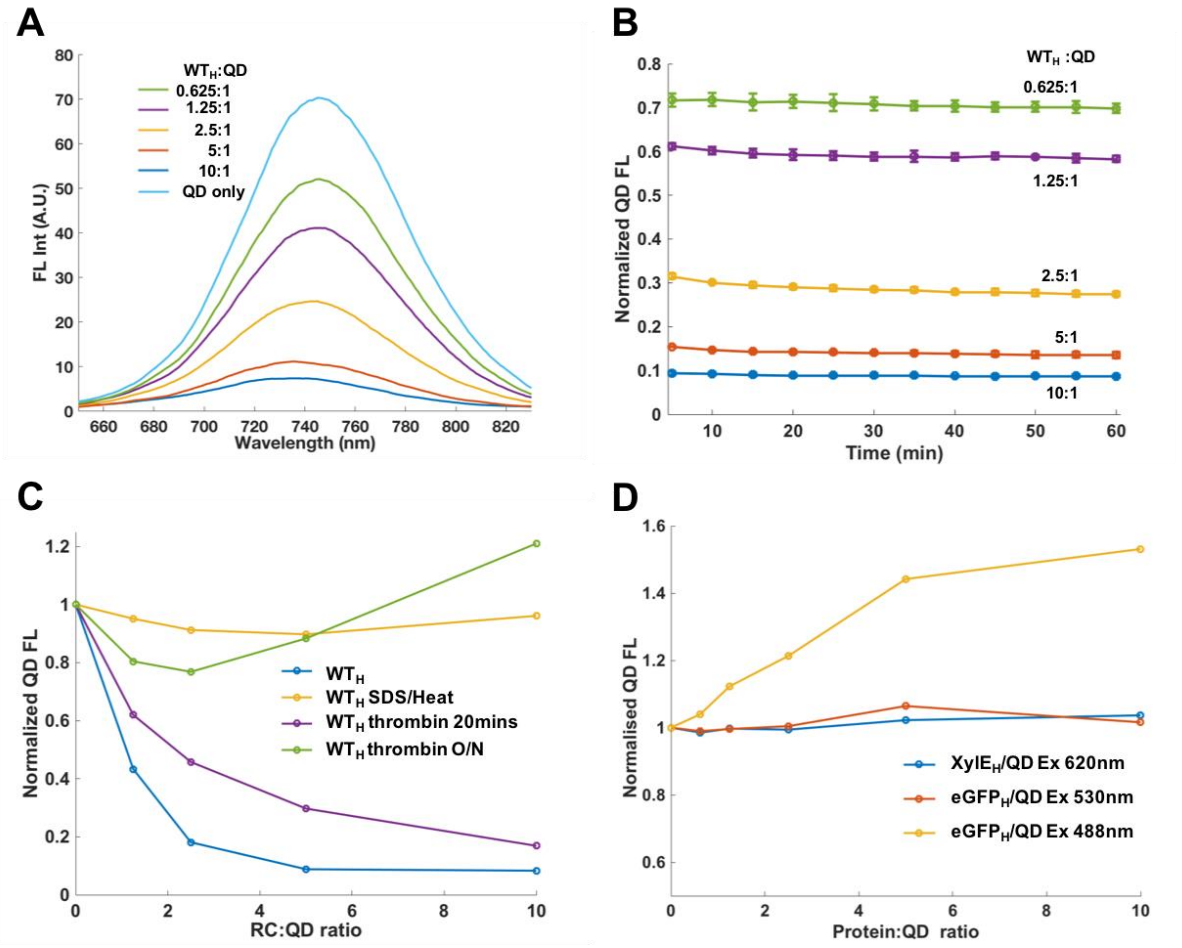


Figure 3.3. Quenching of QD emission by His-tagged WT RCs. (A) Quenching of QD emission by WT_H RCs at different RC:QD ratios. (B) Stability of QD emission quenching by WT_H RCs over one hour at room temperature. (C) Effects of thrombin or SDS/heat treatment on quenching of QD emission by WT_H RCs. (D) Effects of His-tagged XylE or eGFP on QD emission following excitation at the indicated wavelength. Exciting eGFP at 488 nm caused a concentration-dependent increase of QD emission, an effect attributed to energy transfer from the bound eGFP to the QD. In (A)-(C) excitation was at 430 nm where QD absorbance is high and RC absorbance is low and exhibits a trough.

The finding that quenching of QD emission could be reversed by SDS/heat treatment of RC/QD conjugates (**Fig. 3.3C**) indicated that this quenching required the presence of intact RCs, and further suggested that it was not due to a change in the surface chemistry of the QDs caused by binding of the His-tag. To investigate this latter point further, QDs were incubated with two other His-tagged proteins, $XylE_H$, a xylose:proton symport membrane protein, and water-soluble $eGFP_H$. Purification of these proteins is described in Sections 2.7 and 2.8, respectively. Neither produced a decrease in QD emission up to a 10:1 protein:QD ratio (**Fig. 3.3D**).

In addition to decreasing in intensity, it was noticeable that the maximum of the QD emission band blue-shifted somewhat as the RC:QD ratio increased (by a maximum of ~10 nm) (**Fig. 3.3A**). As preparations of QDs are expected to have a distribution of diameters around a certain mean, and hence a distribution of individual emission maxima, a possible explanation is that larger, "red-most" QDs are more effectively quenched by bound RCs due to a better spectral overlap with the RC absorption spectrum. Such a conclusion would be in accord with quenching being due to FRET between the 750 nm emitting QDs and the near-IR absorbing RCs.

3.2.3. Conjugation to QDs enhances RC charge separation.

Proof of FRET between a donor and acceptor typically comprises decreased donor emission accompanied by increased acceptor emission, but in the present case the WT_H RC acceptor is barely fluorescent because excited state energy produces charge separation within a few picoseconds of arrival in the RC. This changes the absorbance spectrum of the RC however, and so it is possible to provide evidence in support of FRET from a QD to a RC by detecting charge separation enhancement in the RC when the attached QD is excited.

Absorbance spectroscopy was used to monitor changes in RC absorbance at 870 nm, associated with photo-oxidation of the P BChls, in response to QD excitation (see Section 2.17). Excitation at 450 nm was used as the WT RC has a relatively low absorbance at this wavelength whereas the absorbance of QDs is relatively high (**Fig. 3.1C**). To further reduce RC absorbance at this wavelength an engineered RC with a glycine to leucine replacement at residue 71 of the M-polypeptide (denoted GM71L) was also used[201]. This structural change prevents incorporation of the single RC carotenoid cofactor (see **Fig. 3.1B**), markedly lowering RC absorbance at 450 nm due to loss of the broad Crt absorbance band between 420 and 580 nm (**Fig. 3.4A**), but does not affect RC charge separation[201].

Photobleaching of P at 870 nm was studied in WT and GM71L RCs with and without a His-tag and in the presence and absence of QDs (RC:QD ratio of 2.5). The procedure is described in Section 2.17. RCs were incubated with a 10-fold excess of UQ_0 to reconstitute the Q_B site and were excited for seven seconds through a 50 nm wide band pass filter centred at 450 nm. Averaged traces were fitted assuming a simple P/P^+ interconversion (see **Eq. 2.1** in Section 2.17). Key fits are compared in **Fig. 3.4B**, with individual data and fits in **Fig. 3.5**. In the absence of QDs the GM71L RCs showed less photobleaching than WT RCs due to their lower absorbance at 450 nm, with no significant differences in the extent of photobleaching between His-tagged and non-His-tagged RCs (**Fig. 3.5**, top half). Addition of QDs to non-His-tagged RCs did not change the extent of photobleaching significantly consistent with a lack of binding (**Fig. 3.5**, right column).

In contrast to findings with RCs without a His-tag, adding QDs to His-tagged WT_H or $GM71L_H$ RCs produced marked increases in photobleaching relative to an equivalent control sample of RCs lacking a His-tag. This comprised a 2.4-fold increase for WT_H RCs over WT RCs, and a 3.1-fold increase for $GM71L_H$ RCs over GM71L RCs (**Fig. 3.4B**). This trend was also seen in the rate constants for P photo-oxidation (k_f) deduced from the fits (**Fig. 3.4C).** These data indicated that energy is indeed donated from photo-excited QDs to bound His-tagged RCs and is used for charge separation, the QDs forming a synthetic antenna complex to complement RC light harvesting in a region of weak absorption.

A much stronger bleaching of the RC P band could be induced by white light excitation of conjugates formed from five different ratios of WT_H RCs and QDs. The dark recovery phase following 0.5 s of photoexcitation could be fitted using a single exponential with a lifetime in the region of ~1.1 s that did not vary significantly between the five different mixtures of RCs and QDs and a RC-only control (**Fig. S3.1** and **S3.2**). This showed that the recombination of $P^+Q_B^-$ was not affected by the presence of the QDs, suggesting a lack of electron or hole transfer between the two when the RC is in a metastable charge-separated state.

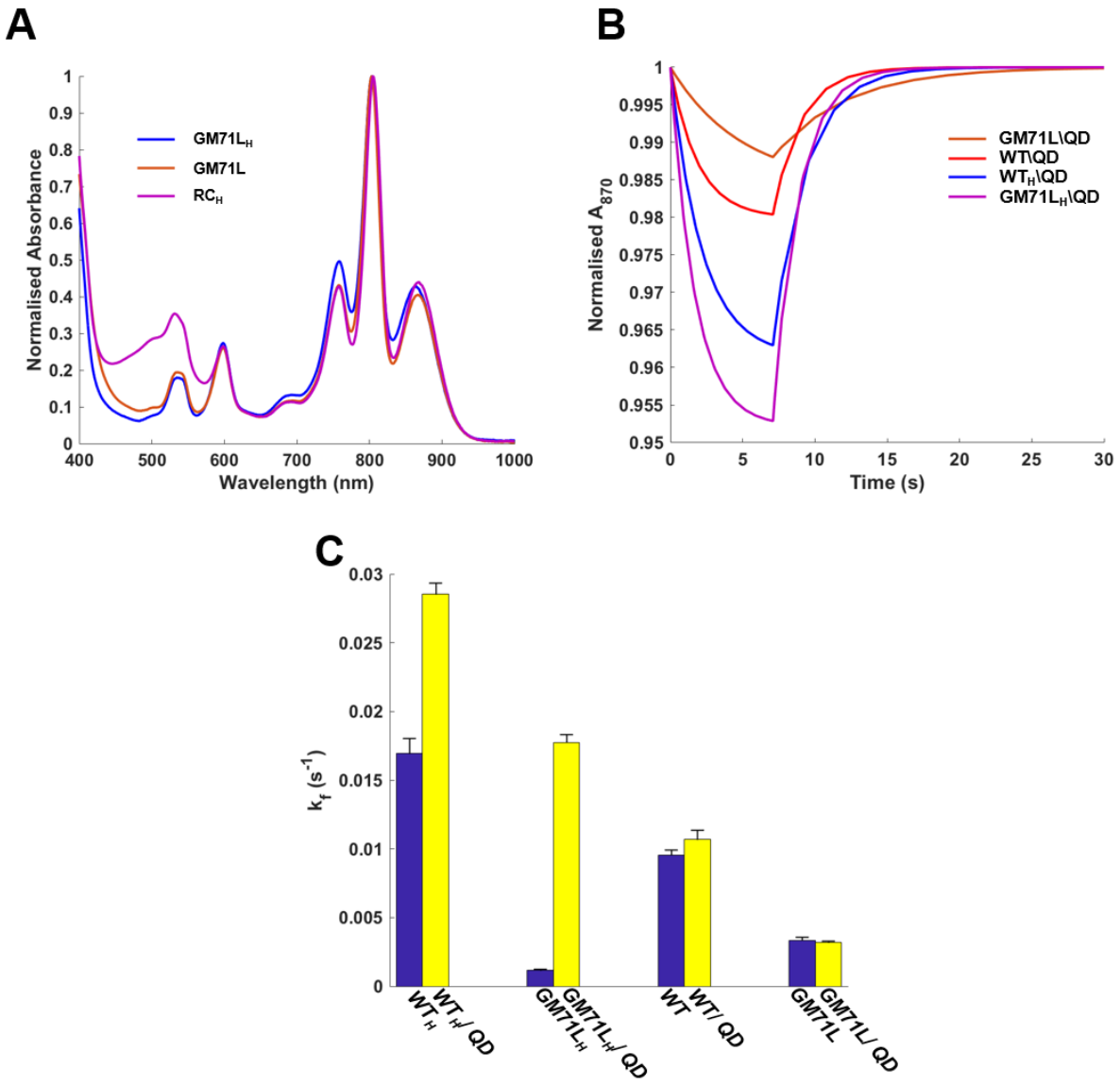


Figure 3.4. Enhanced RC photobleaching in QD conjugates with and without carotenoid. (A) Absorbance spectra. The GM71L and GM71L$_H$ mutant RCs lack the single RC spheroidenone carotenoid which has broad absorbance between 420 and 580 nm. (B) Normalised RC photobleaching at 870 nm in 2.5 RC:QD mixtures of QDs and WT or GM71L RCs with and without a His-tag. Excitation was at 450 nm for 7 seconds. Data shown are fits to averaged kinetic traces (see **Figure 3.5**). (C) Rate constants for photobleaching of P absorbance at 870 nm (k_f). A rate enhancement was seen when WT$_H$ or GM71L$_H$ RCs were conjugated to QDs.

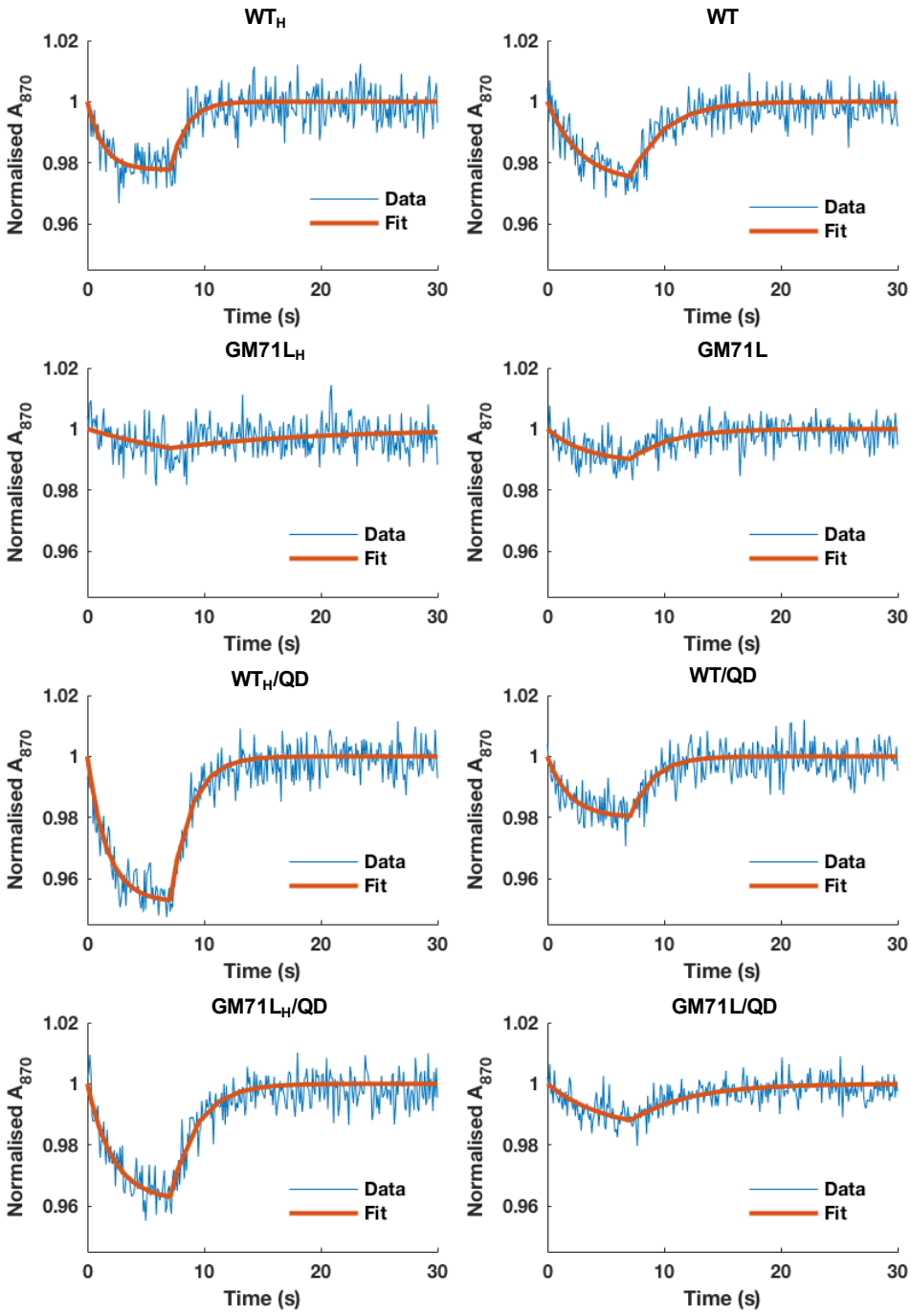


Figure 3.5. RC photobleaching at 870 nm in QD conjugates with and without carotenoid. Each panel shows averaged kinetic traces (n = 5) and a fit using **Eq. 2.1** (section 2.17). The composition of each sample is indicated above each panel. [RC] was 5 μM and [QD] was 2 μM. Excitation was at 450 nm for 7 seconds.

3.2.4. Energy is transferred from QDs to photochemically-inactive RCs.

An alternative approach to demonstrating energy transfer between QDs and bound RCs is to engineer the RC to be fluorescent through of replacement of the valine at position 157 of the L subunit by arginine (VL157R)[202]. This reduces the occupancy of the P dimer binding site to ~0.5 BChls per RC[202], with complete loss of the P absorption band at 870 nm (**Fig. 3.6A**), but has only small effects on the absorbance properties of the remaining monomeric BChl and BPhe cofactors and so the spectral overlap with QD emission. Photo-excitation of these remaining cofactors produces a weak emission band with a maximum at 801 nm, attributable to the monomeric BChls (absorbance band maximum at 798 nm), and a shoulder at 760 nm consistent with a smaller amount of emission from the BPhes (**Fig. 3.6A**, red). Stable charge separation does not occur in the VL157R RC due to the absence of P, measurements by Jackson and co-workers showing that excitation of the monomeric BChls produces emission that decays over a period of several ns, with no indication of the formation of charge separated states such as $B_A^+H_A^-$ or $B_A^+Q_A^-$ [202].

His-tagged VL157R RCs ($VL157R_H$) were prepared by site-directed mutagenesis (Section 2.4). Titration of QD emission using $VL157R_H$ RCs produced a similar quenching curve to that obtained with WT_H RCs (**Fig. 3.6B**, purple), but the spectra included an additional $VL157R_H$ RC emission band (**Fig. S3.3**) that could be deconvoluted (Section 2.16). This demonstrated that quenching of QD emission was independent of the ability of the RC to form metastable cation or anion states such as P^+, H_A^- or Q_A^-. This finding also strongly supports a FRET mechanism for QD quenching as opposed to a mechanism requiring electron transfer to or from the RC.

Comparison of the intensity of weak $VL157R_H$ emission in the absence and presence of QDs showed an interesting effect. As expected, in the absence of QDs the intensity of 801 nm emission from the $VL157R_H$ RC following excitation of its BPhe cofactors at 532 nm was linearly dependent on protein concentration (**Fig. 3.6B**, red)**.** For $VL157R_H$ RC/QD conjugates a greater amount of RC emission was seen at lower RC:QD ratios (**Fig. 3.6B**, cyan compared with red), as might be expected if QDs also excited by the 532 nm light pass energy to bound RCs. However, the difference in $VL157R_H$ RC emission intensity between the conjugate and the corresponding protein-only control was maximal at a RC:QD ratio of 2.5 and then decreased at higher ratios (**Fig. 3.6B**, light orange) despite the extent of quenching of QD emission continuing to increase (**Fig. 3.6B**, purple). The emission from 10:1 conjugates was slightly lower than the equivalent $VL157R_H$ RC-only control (compare red and cyan traces in **Fig. 3.6B**).

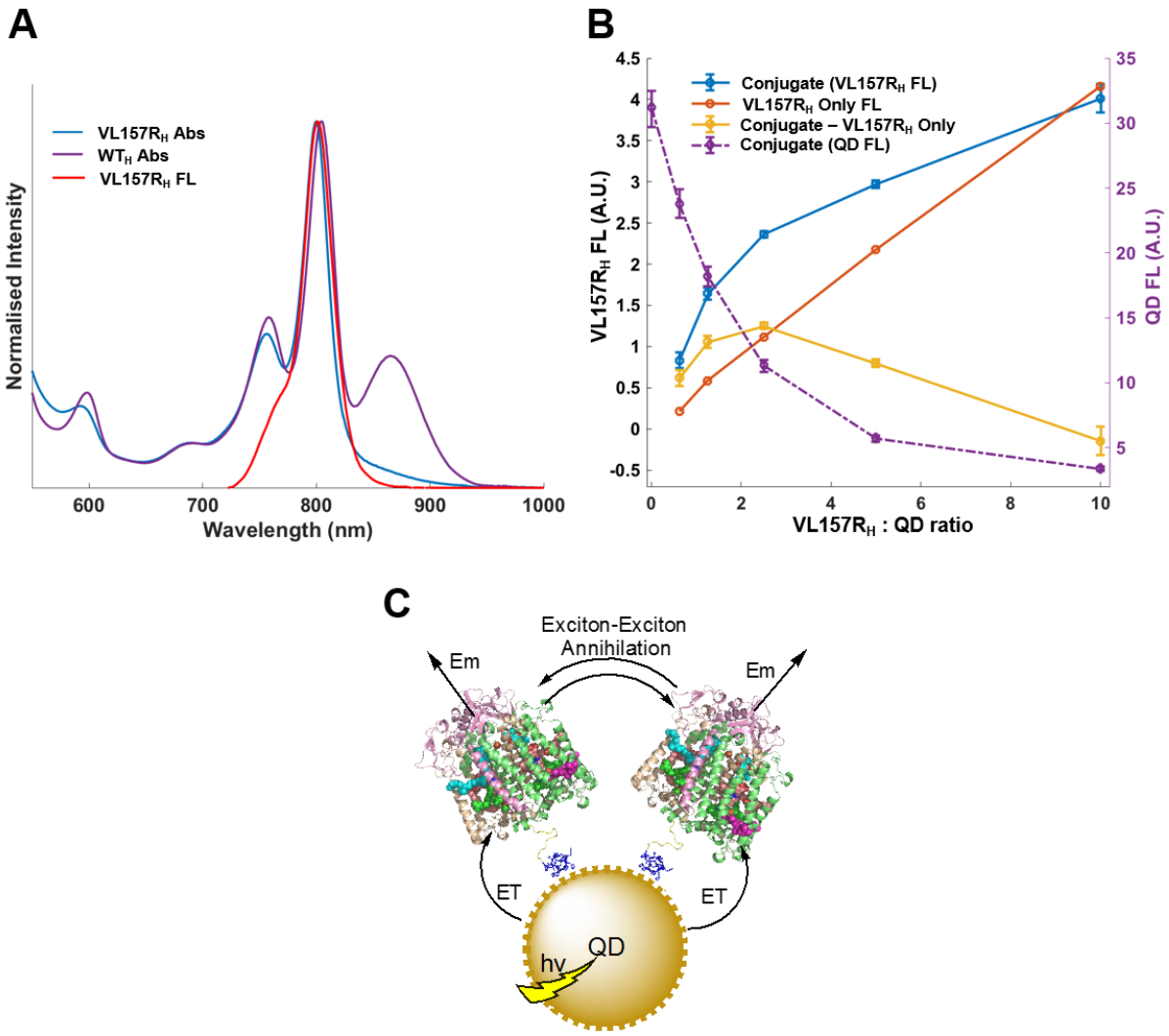


Figure 3.6. Quenching of QD emission by RCs lacking the P BChls. (A) Absorbance and emission spectra for the VL157R_H RC, compared to the absorbance spectrum of the WT_H RC. Spectra are normalised to the signal maximum around 800 nm. (B) Intensity of emission from VL157R_H RCs or QDs at different VL157R_H RC:QD ratios in response to 532 nm excitation. (C) Schematic of the system. Bringing RCs with artificially long-lived excited states into close proximity by linkage to a QD could promote exciton-exciton annihilation, accounting for a decrease in relative emission when packing densities are highest (Technique repeat n = 3).

The decline in VL157R_H fluorescence at higher RC:QD ratios seen in **Fig. 3.6B** (orange) was not due to self-shading, as the intensity of the 800 nm emission from the VL157R_H RCs in conjugates formed in a 10:1 RC:QD mix was linear with the absorbance of the conjugate (**Fig. 3.7A**). In addition, the effect was seen across a range of conjugate concentrations and was linear with concentration (**Fig. 3.7B**). The absorbance spectrum of the VL157R_H RC showed no variance across the range of conjugates formed (**Fig. 3.7CD**), indicating that this effect was unlikely to be due to structural changes to the VL157R_H RC itself.

A feasible explanation for this effect is cross-relaxation between excited VL157R_H RCs as the number bound to each QD increases. Within the nanosecond scale life-time of the excited state formed by direct excitation of the VL157R_H RC, or by energy transfer from a QD (10% ~ 20% of the population was excited considering one exciton generated per absorbed photon), there is the possibility that an exciton can transfer between RCs and relax via a non-radiative pathway such as singlet-singlet annihilation (**Fig. 3.6C**). As such a mechanism requires proximity of excitons, increasing the packing of RCs around each QD makes annihilation more likely, such that the relative amount of emission declines. This effect was not seen for equivalent concentrations of VL157R_H RCs in the absence of QDs because they were free to diffuse independently in solution.

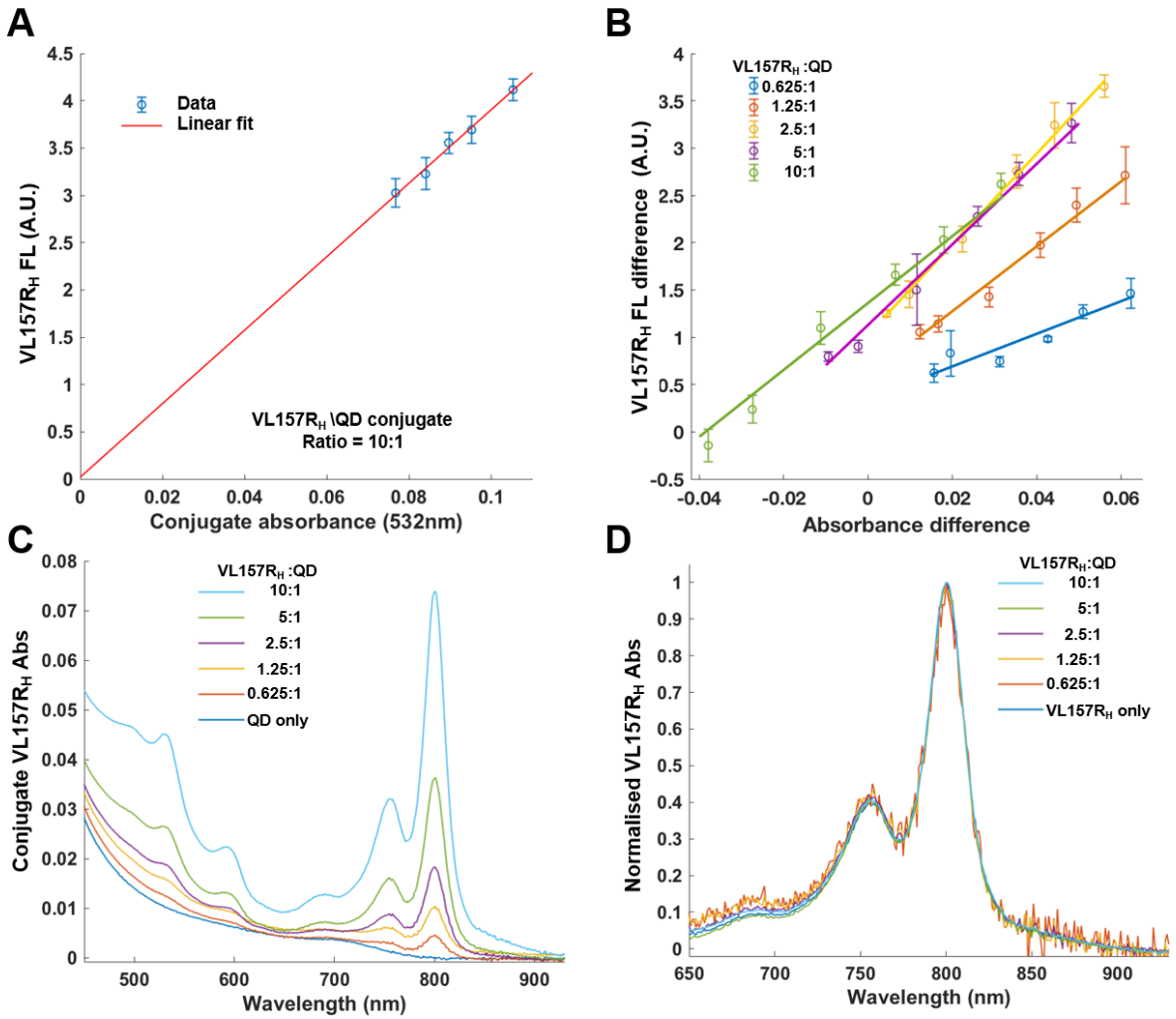


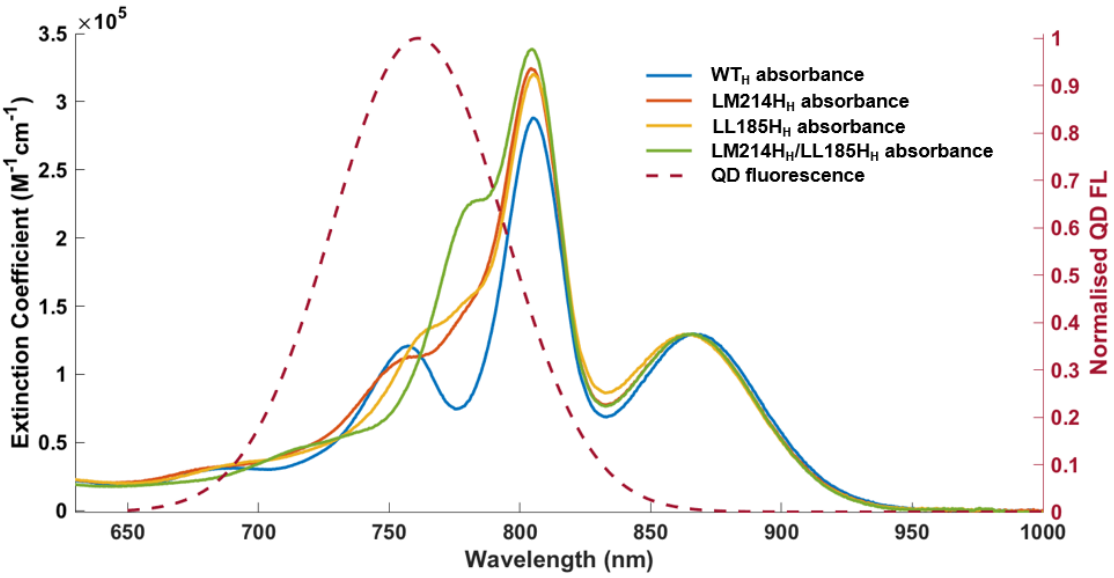
Figure 3.7. Linearity of conjugate emission with concentration. (A) Intensity of 801 nm fluorescence from five different concentrations of VL157R$_H$/QD conjugates formed in a 10:1 RC:QD mix as a function of their absorbance at the excitation wavelength (532 nm), with the line of best fit. The linear dependence demonstrates that the measurements were not affected by self-shading. (B) Additional emission at 801 nm obtained with VL157R$_H$/QD conjugates relative to VL157R$_H$ RCs (532 nm excitation) as a function of additional absorbance of VL157R$_H$/QD conjugates relative to VL157R$_H$ RCs at 532 nm, at 6 sample concentrations. The amount of emission initially rose as the ratio of VL157R$_H$ RCs to QDs increased from 0.625:1 to 1.25:1 to 2.5:1. Beyond this ratio the level of emission dropped for a 5:1 ratio and dropped further for a 10:1 ratio. This pattern was independent of concentration, demonstrating that the fluorescence drop at the higher RC:QD ratios was not due to self-shading effects in the most strongly pigmented samples. (C) Absorbance of VL157R$_H$/QD conjugates. (D) Spectra from (C) with QD absorbance subtracted and then normalized to the Q_y band of B_A/B_B.

3.2.5. QD emission quenching is sensitive to spectral overlap.

Protein engineering can also be used to change the cofactor composition of the RC in a way that does not abolish photochemical charge separation. Mutations LM214H and LL185H are known to cause replacement of a BPhe by a BChl at the H_A or H_B cofactor binding site, respectively,[203,204] whilst the double mutation LM214H/LL185H causes both BPhes to be replaced by BChls. This cofactor change causes the Q_y absorbance band of the H_A and/or H_B cofactor to shift to longer wavelengths (**Fig. 3.8A**), changing the profile of the region of spectral overlap between the RC and QD. The LM214H mutation slows the rate of charge separation by around two-fold and reduces its quantum yield to ~60 % due to replacement of the H_A BPhe by BChl.[203] In contrast the symmetrical LL185H mutation has no discernible effect on charge separation.[204] His-tagged versions of these

His-tagged versions of these three cofactor-exchange RCs were prepared and extinction coefficients were deduced by normalizing their absorption spectra to that of the WT_H RC using the P Q_y band in fully reduced RCs, as the intensity and shape of this band was interfered with least by the mutations (**Fig. 3.8A**). The calculated spectral overlap was increased by between 20 % and 32 % (**Fig. 3.8B)**, and in accord with this the three mutants exerted slightly greater quenching of QD emission than the WT_H RC across the range of RC:QD ratios (**Fig. 3.8C**). This sensitivity to spectral overlap was further evidence for the mechanism of QD emission quenching being FRET to the RC.

A



B

RC_H	Spectral overlap J (cm^3M^{-1})
WT_H	3.780×10^{-12}
$LM214H_H$	4.573×10^{-12}
$LL185H_H$	4.546×10^{-12}
$LM214H/LL185H_H$	4.975×10^{-12}

C

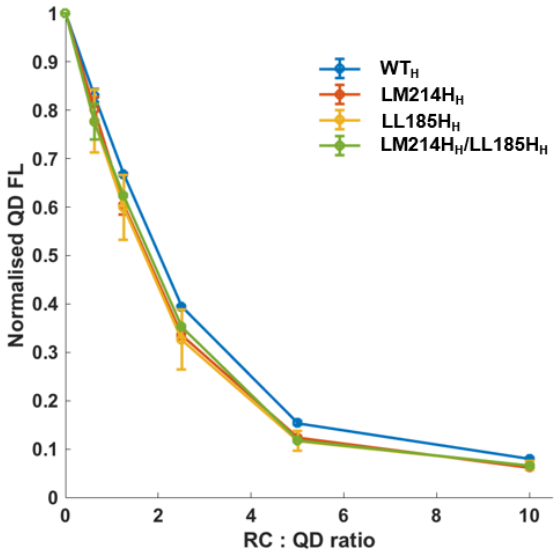


Figure 3.8. Quenching of QD emission by cofactor-exchanged RCs. (A) Absorbance spectra of three RCs with single or double BPhe to BChl replacements compared with that of the WT_H RC and the QD emission spectrum. (B) Calculated spectral overlap. (C) Quenching of QD emission at 750 nm as a function of RC:QD ratio (excitation at 430 nm).

3.2.6. The morphology of RC/QD conjugates.

To investigate morphology, 10:1 RC/QD conjugates constructed with WT_H RCs were examined by TEM (Section 2.12.1). Fast removal of buffer and transfer kept the grid wet to mitigate against drying-induced aggregation prior to fixing with 3% uranyl acetate, the aim being to obtain images that were close representations of the state of the conjugates in solution. Images revealed mostly lightly-stained objects that were evenly distributed in the field of view (**Fig. 3.9A**), suggesting RC/QD conjugates were present as dispersed objects in solution. Multiple images (**Fig. S3.4**) were analysed as described in Section 2.13.2, and compiled data on object diameter were fitted with a lognormal distribution with a mode at ~21.4 nm **(Fig. 3.9B**). This was a physically realistic dimension for an object comprising a shell of RCs surrounding a central QD. Because each RC was connected to its His-tag by a 16 amino acid flexible linker, the maximal theoretical distance between the surface of a QD and a bound RC should be around 6.1 nm, assuming that a fully extended polypeptide chain has a length of 0.38 nm per amino acid[205]. Therefore, the physically plausible range of diameters for a conjugate formed by a QD and multiple RCs should be between a maximum of ~32.7 nm and a minimum of ~20.5 nm based on expected mean diameters for QDs and RCs of 6.5 nm and 7.0 nm, respectively. The fitted distribution was at the low end of this theoretical range (**Fig. 3.9B**), suggesting a closely packed structure.

Conjugate morphology was also examined by DLS (Section 2.13). Scattering profiles did not provide evidence of significant amounts of large-scale aggregate, supporting the conclusion from TEM that conjugates were largely dispersed in solution. The hydrodynamic diameter for 10:1 RC/QD conjugates was 28.4 ± 0.7 nm (**Fig. 3.9C**), again within the theoretical range but larger than the estimate from TEM. The hydrodynamic diameter generally decreased at solution RC:QD ratios below 5 indicating a transition in RC/QD conjugate population from particles with multiple RCs to particles that differed in their diameter by the height of one protein. A difference between mean diameters for 10:1 RC/QD conjugates derived from DLS and TEM was not surprising, as DLS measures the diffusion coefficient of particles and hydrodynamic diameters are generally larger than actual size due to the presence of a solvent shell that migrates with the particle. In contrast negative stain TEM images particles without a water shell adsorbed on a surface. In the present case transient drying during sample preparation could induce RCs to collapse onto the QDs, leading to an underestimation of the true particle size in solution, whereas the DLS data represent an overestimate of the true particle size due to the effect of the water shell.

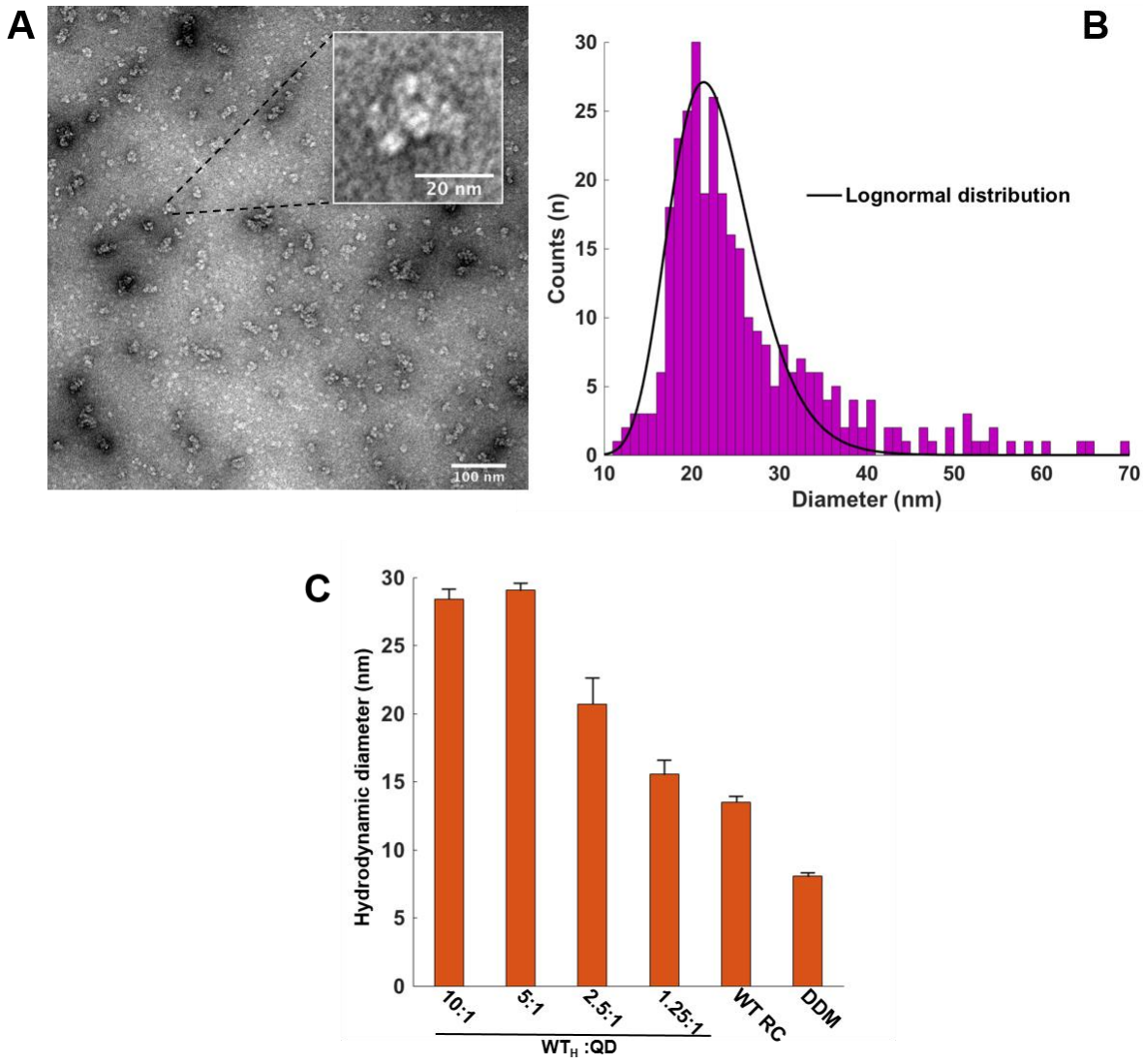


Figure 3.9. Dimensions of RC/QD conjugates. (A) TEM image of RC/QD conjugates formed in 10:1 RC:QD mixture. The inset shows an enlarged view of a typical object, showing multiple clustered bright features of a dimension consistent with the RC protein. (B) Histogram of object diameter for multiple objects identified by image processing, fitted to a lognormal distribution (mode value 21.4 nm). (C) Hydrodynamic diameters of RC/QD conjugates, WT_H RCs and DDM micelles estimated by DLS (average of 5 measurements with standard deviation).

3.2.7. Heterogeneity of actual RC:QD stoichiometry.

For any given mixture of His-tagged RCs and QDs, heterogeneity is expected in the actual RC:QD ratio for individual conjugates because multiple RCs can be accommodated by each QD and binding is a random process. As an understanding of this heterogeneity was important for any estimation of the efficiency of FRET, the interaction between His-tagged RCs and QDs was simulated using a model in which proteins can bind to multiple equivalent and independent sites on the surface of a QD. The model is described by **Eq. 3.1**:

$$\bar{v} = \frac{\bar{n}k_{micro}([P_{free}] - [P_{damage}])}{1 + k_{micro}([P_{free}] - [P_{damage}])} \quad (\mathbf{3.1})$$

where *P* stands for protein and in this Chapter refers to the RC. Parameter $\bar{n}$ is the mean of the maximum number of RCs that can bind to a QD in any mixture (as this maximum number is actually a distribution of values rather than a fixed integer, as there are no discrete binding sites on the surface of a QD). Parameter k_{micro} is the microscopic association constant, and $\tilde{v}$ is the measured average RC:QD stoichiometry in any mixture.

To apply **Eq. 3.1**, the average binding ratio for 10:1, 5:1, 2.5:1, 1.25:1 and 0.625:1 mixtures of WT_H RCs and QDs was determined by separating RC/QD conjugates and free RCs by ultracentrifugation on two-step 0%/25%/60% sucrose density gradients (**Fig. 3.10A**), as described in Section 2.15.3. Careful fractionation of these gradients followed by absorbance spectroscopy revealed that almost all of the protein was pulled to the 25%/60% interface due to being bound to a QD. To correct for continuing dissociation of bound RCs from conjugates during migration to the 25%/60% interface, the proportion of RCs bound to QDs in the initial mixture was calculated by summing the RC contents of fractions 1-8 which were 1 mL aliquots from the bottom of the gradient to just below the 0%/25% interface at the top. The (minor) fraction of “free” RCs in the initial mixture remained at the 0%/25% interface and were collected in fractions 9 and 10. SDS-PAGE/Western blot analysis of these fractions from the 0%/25% interface using an anti-His antibody revealed that, except in the 10:1 RC mixture, most of these free RCs lacked a detectable His-tag, presumably due to its loss during purification, storage or due to photodamage during analysis under ambient illumination (**Fig. 3.10B**). The very small number of His-tagged RCs in the free fraction underscored the tight binding interaction enabled by the tag. Therefore, to calculate $\tilde{v}$ for each mixture the concentration of free RCs in the mixture ($[P_{free}]$) was adjusted for the concentration of RCs not bound to the QDs ($[P_{noHis}]$).

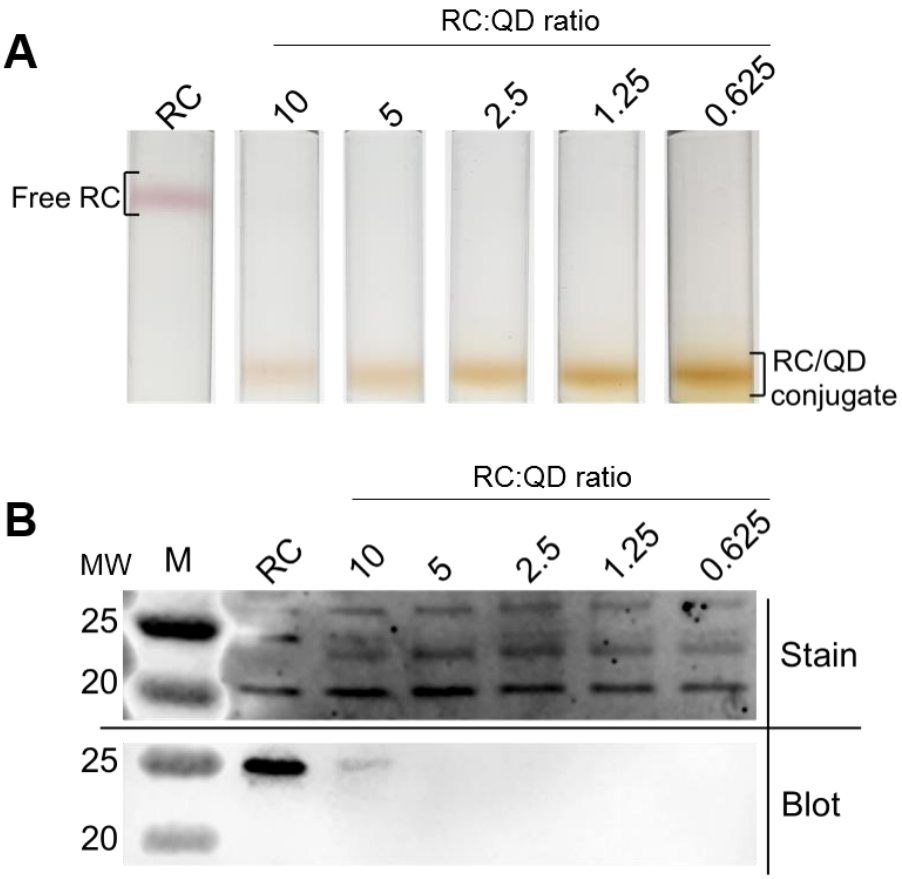


Figure 3.10. Assessment of the fraction of RCs not bound to QDs. (A) Sucrose pull-down experiment in which a fixed concentration of WT_{His} RCs is titrated with increasing concentrations of QDs to achieve the indicated RC:QD ratios. Unbound, free RCs remain that the 0%/25% interface at the top of the gradient and RCs bound to QDs migrate to the 25%/60% interface at the bottom of the gradient. (B) SDS PAGE and Western blot analysis of RCs harvested from the top interface of each gradient. In the stained gel the WT_{His} RC control exhibits three bands corresponding to (top to bottom) the H, M_{His} and L polypeptides of the RC. For the conjugate lanes two bands are obtained for the M polypeptide, an upper band corresponding to His-tagged M and a lower band corresponding to an M polypeptide lacking the His tag. The lower panel shows a Western blot of the same gel using anti-His antibodies. A strong signal is seen for the WT_{His} RC control and a weaker signal in the lane corresponding to a 10:1 RC:QD mix where the fractional population of His-tagged M polypeptide is the greatest. The lower band did not give a signal, confirming that this M polypeptide had lost the engineered His-tag. Protein stain was achieved with Sypro™ Ruby and detected by fluorescence. This was overlaid with A700 channel to present marker together with Sypro™ Ruby signal.

The model summarised by **Eq. 3.1** is shown in diagram form in **Fig. 3.11A**. The fit of this model to the average binding ratio $\tilde{v}$ obtained by applying **Eq. 3.1** to the data from the sucrose gradient pull-down experiments is shown in **Fig. 3.11B**. The parameters $\tilde{n}$ and k_{micro} from the fitting were 15 and 8.23 μM^{-1}, respectively. The value of $[P_{noHis}]$ derived from the fit equated to 4.1 % of the total RC population. The microscopic thermodynamic constant k_{micro} was converted to a macroscopic dissociation constants (K_d) at permitted valencies (i) using **Eq. 3.2:**

$$K_{d,i} = \frac{i}{(\bar{n} - 1 + i)k_{micro}} \quad (i = 1, 2, ..., \bar{n}) \tag{3.2}$$

This gave a high binding affinity with $K_{d,i=1}$ = 8.1 nM between a free QD and the first RC.

Conjugate assembly and disassembly was addressed through the model depicted in **Fig. 3.11a** and the reaction scheme:

$$P_{His} + P_{i-1}QD \underset{k_{r,i}}{\overset{k_{f,i}}{\rightleftharpoons}} P_iQD \quad (i = 1, 2, ..., \bar{n}) \tag{3.3}$$

where P_{His} is the total RC protein population adjusted for RCs without a His-tag and so unable to bind, and $k_{f,i}$ and $k_{r,i}$ are the macroscopic kinetic constants for binding and dissociation, respectively, at permitted valences (i) from 0 to $\tilde{n}$.

A set of ordinary differential equations (ODEs) was then generated at applied RC:QD ratios of 0.625, 1.25, 2.5, 5 and 10, as shown below:

$$\frac{d[P_{His}]}{dt} = \sum_{i=1}^{\bar{n}} (k_{r,i}[P_iQD] - [P_{His}]k_{f,i}[P_{i-1}QD]) \tag{3.4}$$

$$\begin{aligned} \frac{d[P_iQD]}{dt} =& [P_{His}](k_{f,i}[P_{i-1}QD] - k_{f,i+1}[P_iQD]) \\ &+ k_{r,i+1}[P_{i+1}QD] - k_{r,i}[P_iQD] \\ &(i = 0, 1, 2, ..., \bar{n}) \end{aligned} \tag{3.5}$$

where $k_{f,i}[RC_{i-1}QD]$ at i = 0 and $k_{r,i+1}[RC_{i+1}QD]$ at $i = \tilde{n}$ were omitted.

The macroscopic kinetic constants $k_{f,i}$ and $k_{r,i}$ used in **Eqs. 3.4 and 3.5** were determined from their corresponding microscopic kinetic constants using:

$$k_{f,i} = (\bar{n} + 1 - i)k_{on,micro} \quad (i = 0, 1, 2, ..., \bar{n}) \tag{3.6}$$

$$k_{r,i} = ik_{off,micro} \quad (i = 0, 1, 2, ..., \bar{n}) \tag{3.7}$$

A deterministic simulation was carried out in MATLAB using the ODEs and the defined kinetic constants that enabled an estimate of conjugate heterogeneity for any RC concentration and RC:QD ratio. The results of the simulation process are summarized in **Fig. 3.12**, showing that each system came to a stable equilibrium state from which the final population distribution could be determined.

Fig. 3.11C shows the resulting population distribution for a 10:1 RC:QD mixture ([QD] = 50 nM), which produced a mean of 7.7 RCs per QD and individual stoichiometries ranging from 1:1 to 15:1. Distributions in RC:QD stoichiometry for other overall ratios of RC:QD are summarized in **Fig. 3.13.**

It was previously reported that the stoichiometry of conjugates formed between His-tagged maltose binding protein (MBP) and QDs follows a Poisson distribution[198]. Comparison of the present simulation result to a Poisson distribution with the same mean showed a broadly similar but not identical structure especially at high RC:QD ratios (**Fig. 3.11C**). A more realistic explanation of the observed heterogeneity in RC:QD stoichiometry is that formation of RC_i/QD conjugates (where *i = 1,2,…,ñ*) depends on all previous species whereas a Poisson distribution assumes every event happens independently. This could explain why the Poisson distribution predicts higher population of complexes with very high RC:QD stoichiometries than the simulation.

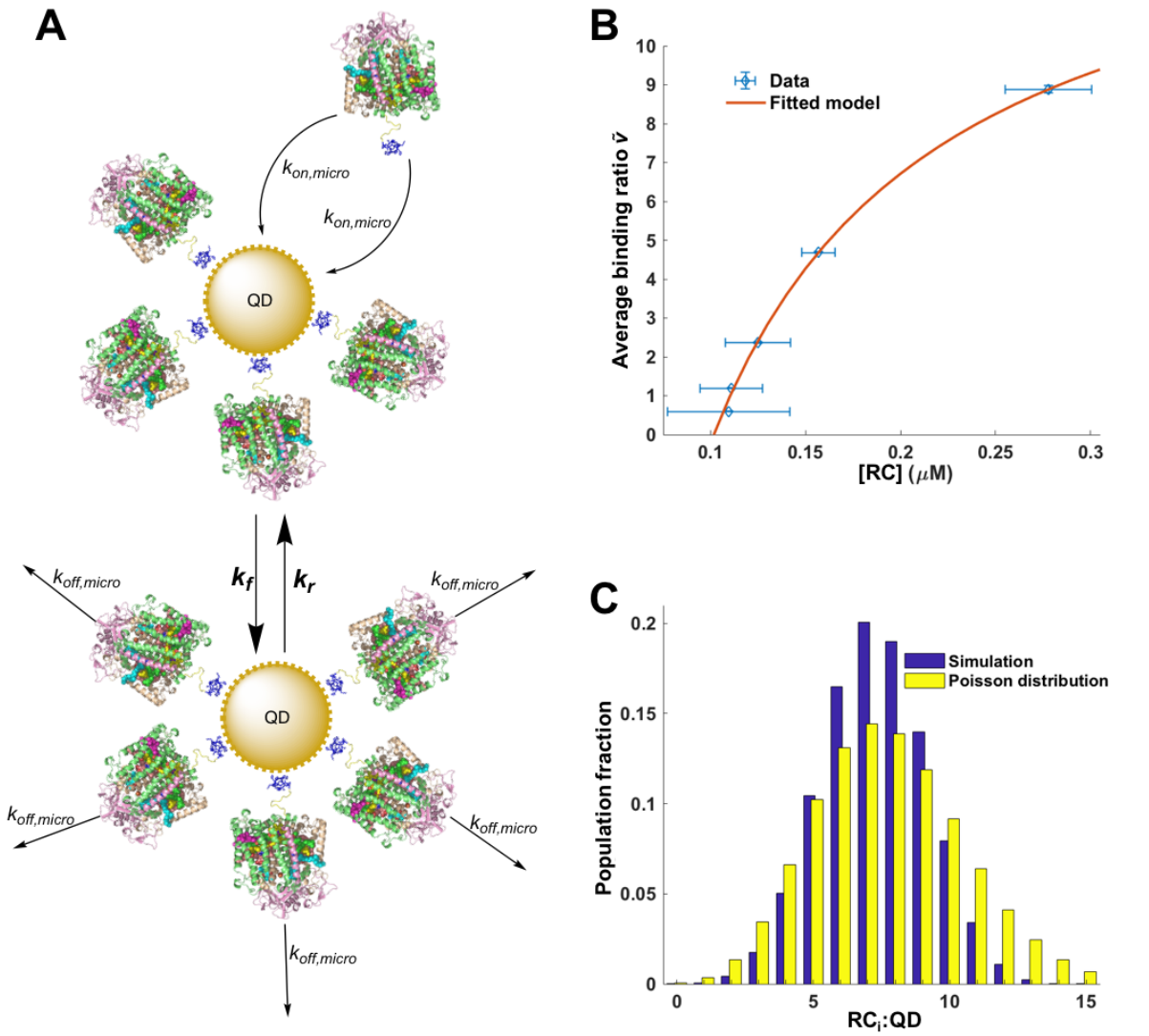


Figure 3.11. Assessment of distribution in RC:QD stoichiometry. (A) Schematic of model for simulation of binding of His-tagged RCs to multiple equivalent and independent sites on the surface of a QD. (B) Data on average binding ratio as a function of the concentration of unbound RCs fitted using **Eq. 3.1.** (C) Histogram of the fraction of the conjugate population with a certain RC:QD ratio for a mix formed from 10 RCs per QD. The distribution peaked at 7.7 RC per QD with a maximum stoichiometry of 15. A calculated Poisson distribution is shown for comparison.

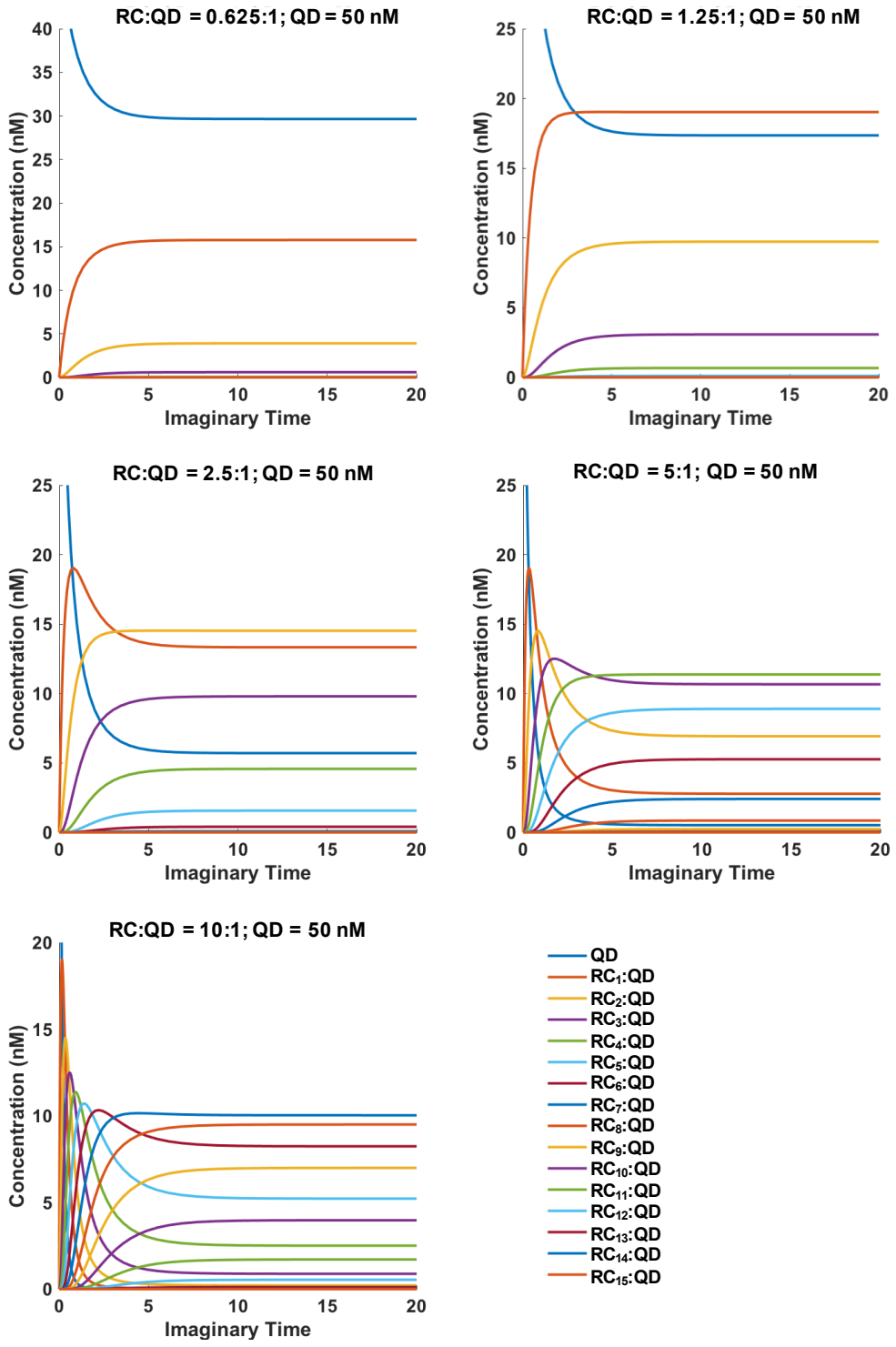


Figure 3.12. Modelling of the distribution in RC:QD stoichiometry. Each panel shows determination of the concentration of RC/QD conjugates with different RC:QD ratios within a population formed from mixes with different stoichiometries of RC to QD. The key to all panels is shown in the bottom right.

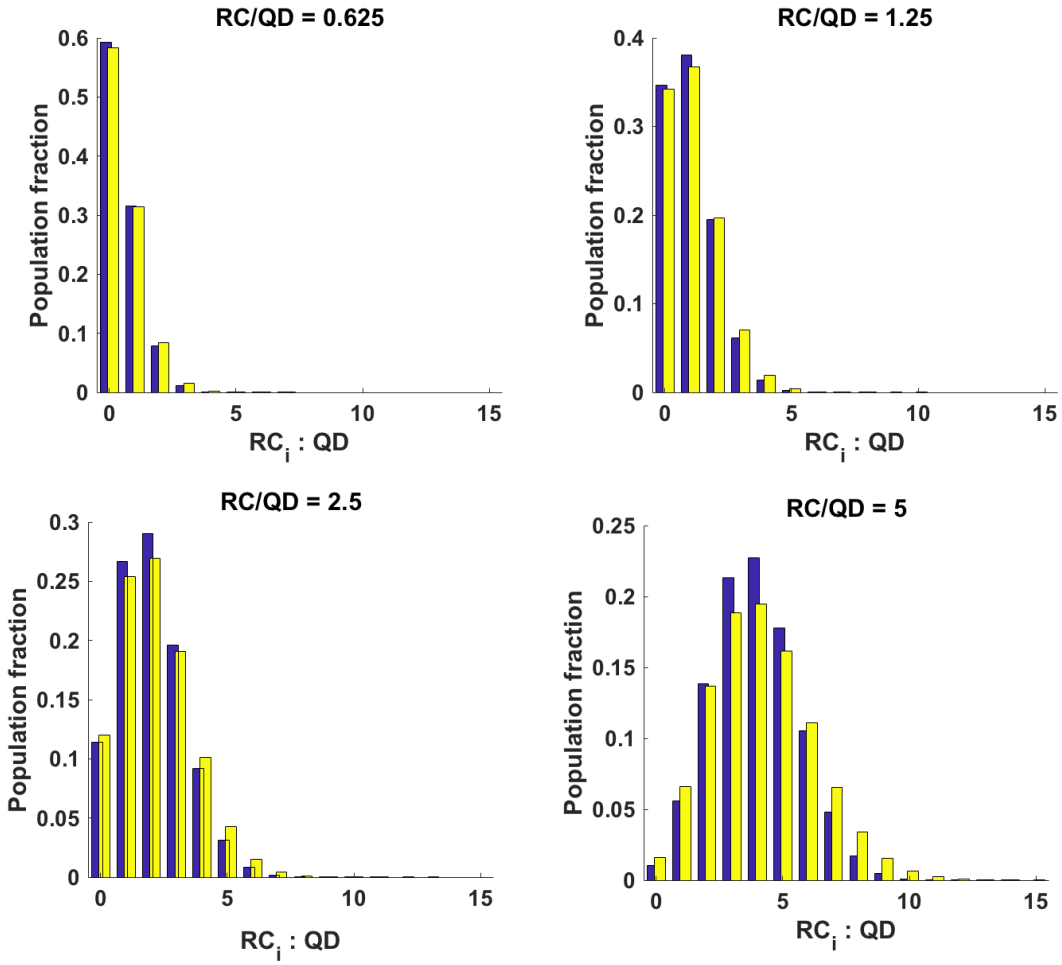


Figure 3.13. Distributions in RC:QD stoichiometry at different overall RC:QD ratios. Each panel shows a histogram of the fraction of the conjugate population with a certain RC:QD ratio for four mixes formed from between 0.625 and 5 RCs per QD. In each case the modelled distribution (blue) is compared to a Poisson distribution (yellow). Data for RC/QD = 10 are shown in **Fig. 3.12C**.

3.2.8. Estimations of FRET efficiency and distance.

The standard analysis of a FRET interaction considers a donor-acceptor pair, but in the present case, even at RC:QD ratios below unity, an individual QD donor can accommodate multiple RC acceptors. Accordingly, the average energy transfer efficiency between a single QD-RC pair (E_{DA}) was calculated using **Eq. 3.8**, which took into account the heterogeneity in QD:RC stoichiometry. The theoretical correlation between E_{DA} and the apparent efficiency for a multiple acceptor system (E_{app}) is shown in **Fig. 3.14A.** The relationship was close to linear for conjugates where the average RC:QD was below one, and became increasingly non-linear as RC:QD increased. This non-linear relationship underscored the importance of accounting for heterogeneity in seeking a precise understanding of the efficiency of FRET from a QD to multiple bound RC acceptors. Based on these above considerations, single QD-RC pair FRET efficiencies (E_{DA}) at different RC:QD ratios were estimated using **Eq. 3.8.**

$$E_{app} = \sum_{i=1}^{\bar{n}} p(i) \frac{iE_{DA}(i)}{(i-1)E_{DA}+1} \qquad \textbf{(3.8)}$$

The term $p(i)$ where (i = 1,2,…,$\tilde{n}$) represented the distribution of conjugates with different numbers of RC per QD. Values of p were determined from the simulation described above in Section 3.2.7.

Values of E_{DA} increased as RC:QD increased and plateaued above a 5:1 ratio at a value of around 0.64 for conjugates formed from WT RCs (**Fig. 3.14B**). Uniformly higher FRET efficiencies were obtained for the three BPhe replacement mutants for every assessed RC:QD ratio (**Fig. 3.14B**), again maximising at a RC:QD of 5:1. This higher efficiency likely arose from the greater spectral overlap in these cofactor-exchanged RCs (see above, Section 3.2.5).

The distance for 50 % FRET efficiency, R_0, was determined from **Eq. 3.9**[197]:

$$R_o^6 = 8.785 \times 10^{-25} \frac{\Phi_D \kappa^2 J}{n^4} \quad (in\ cm^6) \qquad \textbf{(3.9)}$$

where Φ_D is the QD quantum yield, J is the spectral overlap between donor fluorescence and acceptor absorbance, and n is the refractive index of water (1.33). The orientation factor κ^2 was assumed to be ⅔. The quantum yield of the QDs was determined as described in Section 2.11.1,

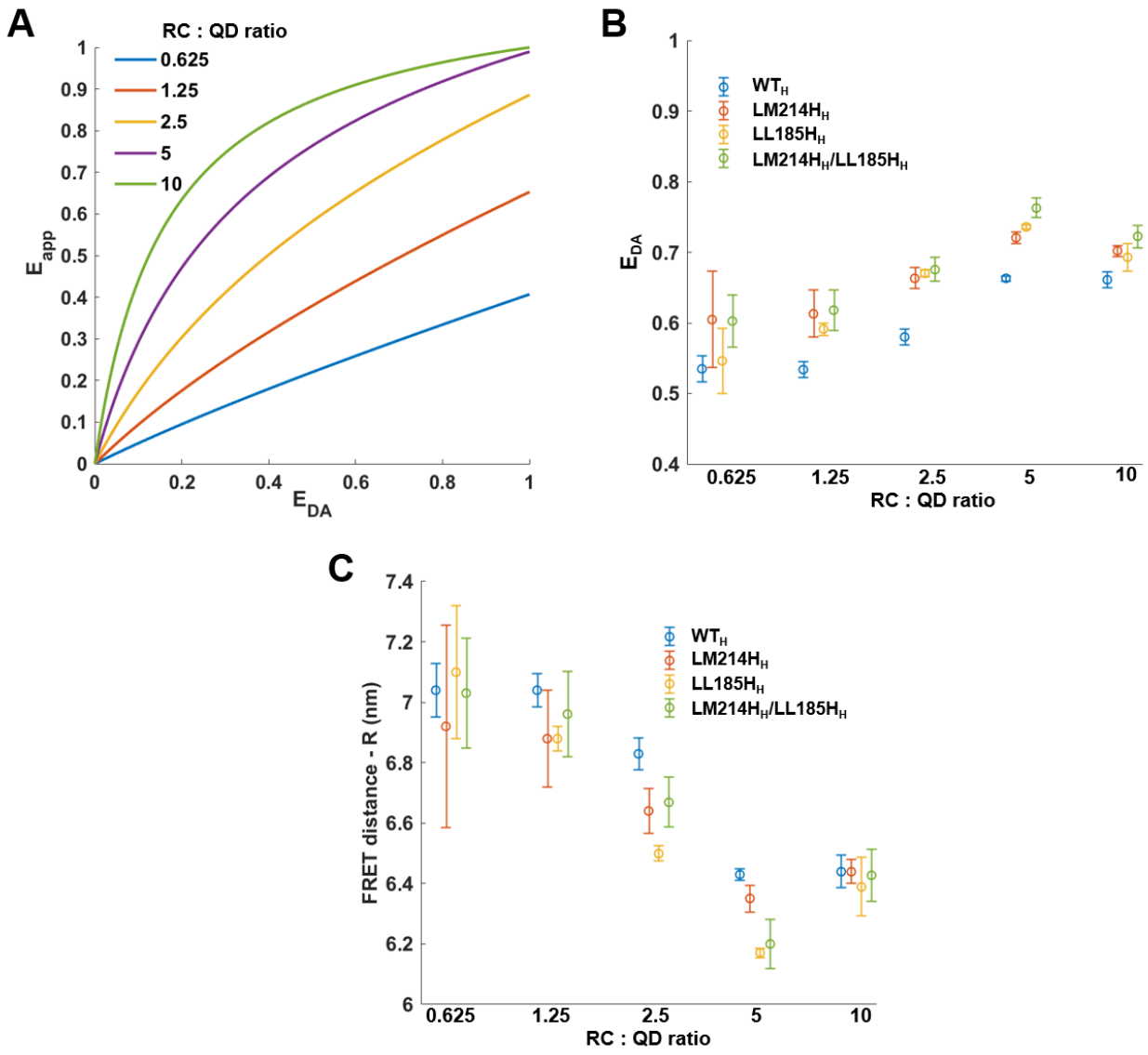
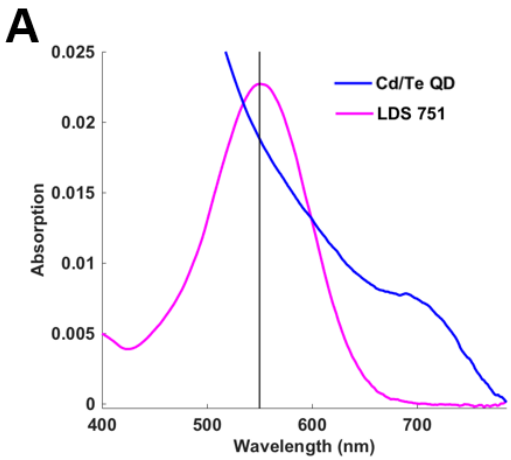
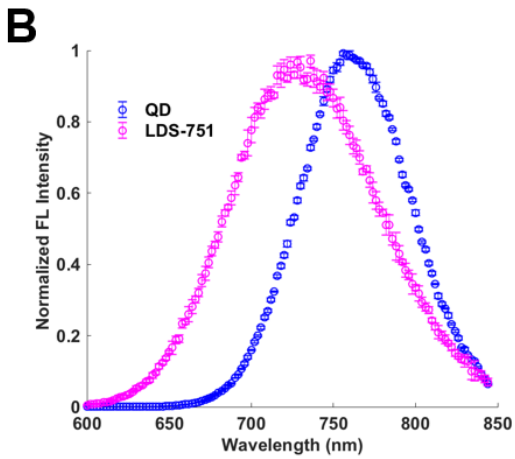


Figure 3.14. Estimations of FRET efficiency and distance. (A) Calculated correlations between the apparent FRET efficiency and the FRET efficiency for a single donor-acceptor pair for different ratios of RC:QD determined using **Eq. 3.8**. (B) Calculated FRET efficiency for a single donor-acceptor pair as a function of RC:QD ratio determined from E_{app} and the correlation in (A). (C) Calculated FRET distance (R) as a function of RC:QD ratio, determined from **Eqs. 3.9** and **3.10**.





C

$XylE_H$:QD	Quantum Yield
10	0.196 ± 0.001
5	0.195 ± 0.002
2.5	0.202 ± 0.002
1.25	0.208 ± 0.002
0.625	0.188 ± 0.002
0	0.192 ± 0.002
Mean	0.197 ± 0.007

Figure 3.15. Determination of quantum yield of QDs. (A) Absorbance spectra of water-soluble, 6.5 nm diameter CdTe QDs in 20 mM Tris (pH 8)/0.04 % DDM and the dye LDS-751 in methanol. (B) Averaged emission spectra (550 nm excitation, 10 spectra). (C) Quantum yields for QDs or $XylE_H$/QD conjugates formed from mixes with different $XylE_H$:QD ratios. The procedure for calculating quantum yields is described in Section 2.11.1.

using the data summarised in **Fig. 3.15**. This involved the use of the optically-inactive His-tagged membrane protein $XylE_H$ to simulate generic protein binding to QDs.

Having determined E_{DA} and R_0, the actual FRET distance, R, was then estimated using **Eq. 3.10**:

$$R^6 = \frac{1 - E_{DA}}{E_{DA}} R_o^6 \qquad (\mathbf{3.10})$$

The values of R arrived at for the WT_H RC and cofactor replacement mutants in five different conjugates are shown in **Fig. 3.14C**. These values were physically reasonable, for WT_H varying between ~7.0 nm for the lowest RC:QD ratio and dropping to ~6.4 nm for the highest RC:QD ratio (**Fig. 3.14C**). The calculated conjugate diameter based on the FRET distances was therefore 22.1 $\pm$ 0.2 nm and 20.9 $\pm$ 0.1 nm from the lowest and highest RC:QD ratios (assuming the FRET distance is from the centre of the QD to the centre of the four acceptor RC bacteriochlorins). These deduced conjugate sizes were close to the estimated diameter equal to 21.4 nm from TEM.

Consistency in the calculated FRET distances for the WT_H RC and three BPhe replacement mutant RCs suggested the generally improved QD emission quenching seen with the latter is indeed due to their altered absorption spectra and improved spectral overlap with the QD emission.

3.3. Discussion

As outlined in Chapter 1, there is growing interest in the use of proteins from the photosystems of *Rba. sphaeroides* and related anoxygenic purple photosynthetic bacteria in devices that take advantage of their highly quantum-efficient light harvesting and separation of electrical charge[91,106,109]. A drawback, however, is their limited energy harvesting across a large part of the visible region of the solar spectrum.

One option to address this is to artificially augment the light harvesting capacity of the *Rba. sphaeroides* RC (see review[206]). Published approaches include the direct attachment of blue-absorbing or green/red-absorbing synthetic dyes to lysine residues on a carotenoid-less RC,[207,208] the direct attachment of green/red or red/nearIR-absorbing dyes to cysteine residues engineered on native RCs[209] and the attachment of pairs of dyes carried on a DNA nanoscaffold.[210] These provide a means of adding multiple chromophores to a RC, although it is difficult to control their spatial arrangement when multiple attachment sites are present. Augmented energy harvesting has also been studied in a fusion protein between a carotenoid-deficient RC and a yellow fluorescent protein.[133]

Three previous experimental studies have looked at energy transfer between purified *Rba. sphaeroides* RCs and water-soluble QDs[211–213]. All three of these somewhat related studies employed WT RCs without a His-tag that had been purified from a native strain of *Rba. sphaeroides* through LDAO-solubilisation and hydroxyapatite column chromatography. Where discussed, the binding of RCs to QDs inferred from quenching of QD emission was attributed to unspecified electrostatic interactions. In contrast, in the present work a His-tag was used to bind RCs to QDs with nanomolar affinity for the first interaction and in an oriented fashion with the bacteriochlorin cofactors closest to the QD. This tight and specific binding allowed conjugate formation at a QD concentration of 50 nM, some 5-20 fold lower than in previous work[211–213], avoiding the need to correct data for inner filter and reabsorption effects arising from carrying out fluorescence spectroscopy on micromolar concentrations of RC ([RC] was maximally 500 nM in the present work). His-tagged RCs were also purified from strains of *Rba. sphaeroides* lacking the genes that encode the light harvesting complexes,[86] preventing contamination of RCs with LH2 complexes as reported previously[211]. Site-directed mutagenesis was also used to remove the RC carotenoid to enable more selective QD excitation, to vary donor-acceptor spectral overlap by replacing BPhe cofactors with BChls, and to render the RC fluorescent to confirm the energetic link between RCs and QDs.

The His-tag used to anchor the RC to the surface of a QD was genetically engineered to the protein by a linker sequence that can be modified at will. As the dimensions of the QDs

and RCs are known, and the maximum length of the linker can be estimated, it is possible to deduce the number of RCs that can pack around a QD at different RC-QD separations dictated by the linker length. This was done by reducing the problem to the number of circles of a diameter of 10 nm, representing the maximum cross-section of a RC-detergent micelle complex, that can fit on a spherical surface of different diameters (**Fig. 3.16**). The permitted number of RC-detergent complexes is 4 at a separation (i.e. linker length) of 0 nm and increased to 17 at a maximal separation of 6.08 nm (**Table 3.1**). This upper packing limit corresponded reasonably well to an upper limit of 15 RCs per QD derived from modelling of the RC-QD interaction (see **Fig. 3.11** and **Fig. 3.13**).

A notable variance from previously published work was the high ratio of RC:QD (10:1) required for high (~92 %) quenching of QD emission in the present study. In the most comparable previous example, Maksimov and co-workers[213] reported a maximum of 85% quenching of 780 nm emission from CdTe QDs at a RC:QD of 3.6. This much lower maximal ratio with similarly sized QDs is consistent with direct electrostatic binding of RCs to the 780 nm QDs that saturates around a ratio of 4 RCs per QD (**Table 3.1**). Previous titrations of smaller QDs with RCs have also reported maximum quenching as occurring at ratios of between 2 and 6 RCs per QD, consistent with direct electrostatic binding of the protein to the QD surface that limits the number of RCs that can pack around a QD. A drawback is that this type of binding interaction cannot be easily manipulated. In contrast, the strategy of using a His-tag and linker in the present work not only facilitated strong and specific binding but also permitted a larger number of RCs to be assembled around each QD due to the details of a programmable linker, maximising the extent of QD emission quenching.

The principal conclusion arrived at in the work described in this chapter is that energy was passed from QDs to bound RCs by FRET, with no indication of energy transfer mediated by an alternative process such as electron exchange. Measurements revealing QD enhancement of RC photobleaching in WT_H and carotenoid-less RCs, and enhancement of BChl emission from P-less RCs, clearly showed that energy is passed from QDs to the RC BPhe and monomeric BChl pigments when a conjugate is formed between multiple proteins and a QD. Quenching was not seen using other His-tagged proteins or following SDS/heat treatment to unfold intact WT_H RCs, or following imidazole/histidine treatment to unbind WT_H RCs, or following protease treatment to detach the WT_H RC from its tag, showing it to be dependent on conjugation to structurally-intact RCs. In contrast, quenching was unaffected by removal of the carotenoid cofactor, replacement of one or both of the RC BPhes with BChl, or removal of the primary donor BChls, showing that it did not involve the carotenoid cofactor or require the operation of conventional photochemical charge separation.

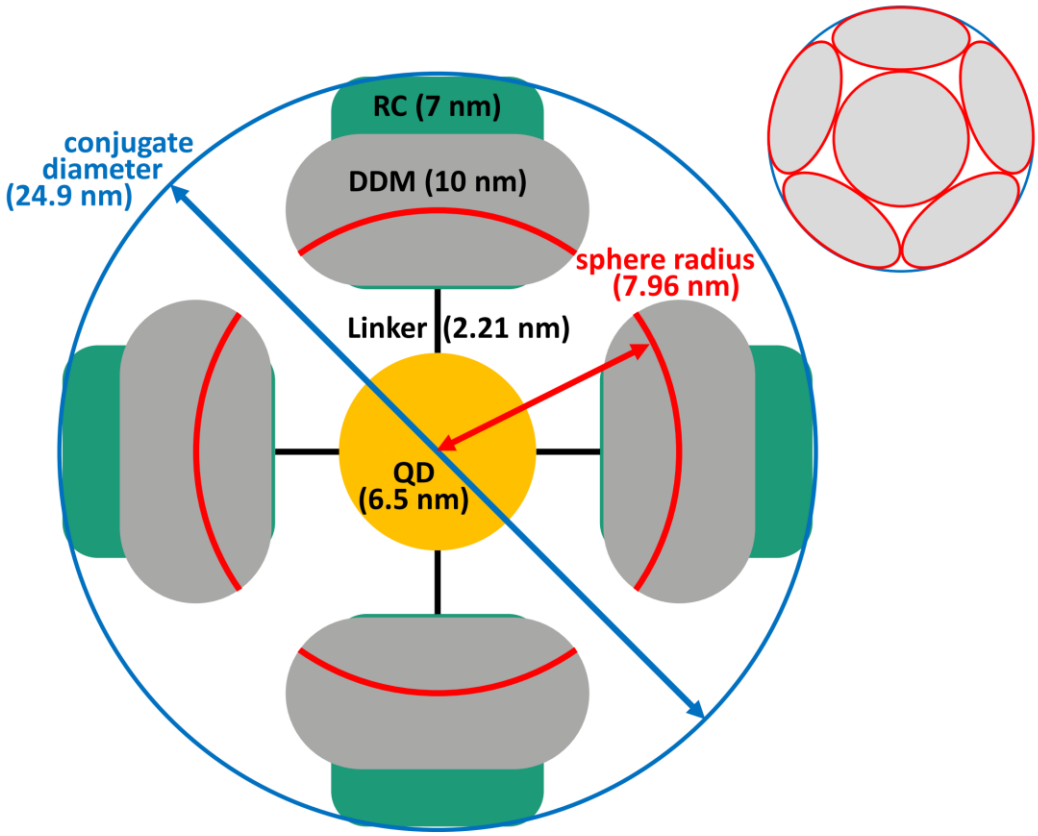


Figure 3.16. Illustration of RC packing around a central QD. In a schematic of the model, RCs (green) in detergent micelles (grey) are attached to the central QD (orange) by a linker (black line) of between 0 and 6.1 nm in length. The length of the linker determines the radius of a sphere (red) along with the radius of the QD (3.25 nm) and a distance of 2.5 nm between the surface of the RC and the mid-point of the thickness of the detergent micelle. The model determines how many circles of radius 10 nm can be fitted on the surface of this sphere without overlapping (see inset), each circle representing a cross-section in the plane of the membrane through a 7 nm diameter RC in a 3 nm diameter DDM micelle[191]. The length of the linker also determined the overall diameter of the conjugate (blue). The specific example shown is for a conjugate with an overall diameter mid-way between experimental estimates from TEM and DLS with a linker of 2.21 nm.

Table 3.1. Effect of linker length on the maximum number of RCs that can pack around a QD and overall conjugate diameter.

Linker[a] length (nm)	Sphere[b] radius (nm)	Max packing[c] number RC/DDM (n)	Max packing[d] number RC/LDAO (n)	Conjugate[e] diameter (nm)	Experimental[f] estimates
0.00	20.50	4	4	20.50	
0.49	21.48	4	6	21.48	TEM - 21.4 nm
0.98	22.47	4	7	22.47	FRET – 22.1 nm
1.47	23.45	6	8	23.45	
1.97	24.43	6	9	24.43	
2.21	24.80	7	10	24.82	
2.52	25.54	8	12	25.54	
3.01	26.52	9	12	26.52	
3.50	27.50	10	14	27.50	
3.99	28.48	12	15	28.48	DLS - 28.4 nm
4.48	29.47	12	17	29.47	
4.97	30.45	13	18	30.45	
5.47	31.43	15	20	31.43	
6.08	32.66	17	24	32.66	

[a] Theoretical length of the linker connecting the body of the RC to its His-tag.

[b] Radius of a sphere determined by the 3.25 nm radius of a QD, the length of the linker, and a distance of 2.5 nm between the surface of the RC and the mid-point of the thickness of the detergent micelle (see **Fig. 3.16**).

[c] Maximum number of non-overlapping circles of 10.2 nm diameter that can be fitted on this sphere.

[d] Maximum number of non-overlapping circles of 8.9 nm diameter that can be fitted on this sphere.

[e] Overall conjugate diameter determined by the diameter of one QD, diameters of two RCs and two linker distances.

[f] Experimental estimates of conjugate diameter from DLS and TEM.

I conclude from this that the basic requirement for quenching of QD emission by RCs is one or more BChl or BPhe pigments with absorbance that overlaps with the QD emission band. Once transferred, energy can be dissipated through charge separation (as in the WT_H and cofactor replacement mutants) or BChl or BPhe emission (as in the $VL157R_H$ P-less RC). An interesting observation made with $VL157R_H$/QD conjugates was that RC emission initially increased as the RC:QD ratio increased up to a value of 2.5:1, as might be expected given the associated increase in quenching of QD emission, but then declined as this ratio increased further to 5:1 and then 10:1 despite the fact that quenching of QD emission continued to increase. The reason for this remains to be confirmed but the most likely is exciton-exciton annihilation. Such annihilation processes are well documented in natural arrays of purple bacterial light harvesting complexes,[214,215] and in recent years have been studied systematically in LH2 arrays reconstituted into artificial bilayers,[216,217] but would not normally be detected in purple bacterial RCs due to efficient energy trapping by charge separation.

Heterogeneity in the number of RCs per QD is a factor that has an influence over the analysis of the efficiency of FRET. In the present work this heterogeneity could be well-explained by thermodynamic scheme based on an independent binding model. Although the QD surface did not offer a distinct number of discrete “binding sites”, the model offered a good approximation of the conjugation process between multiple RCs and a QD. Considering that protein/QD conjugation is the result of a collision between the protein His-tag and the QD surface, when any “binding site” is already occupied by a protein, further attachments to this area are prohibited. Also, because of the dynamics of a protein attached to a site by a flexible linker, the shielding of binding could happen over a rather larger area than that of the protein itself and might be affected by neighbouring proteins. In addition, there is potential for reorganisation of bound proteins on the QD surface. Taking these factors together, it is to be expected that the final number of “binding sites” will be heterogeneous around an average. It has been reported that the kinetics of His-tag mediated binding of proteins to QDs occurs rapidly[218] and in the work described here it was observed that self-assembly was stable within a few minutes of initiation of binding, indicating that the system was at equilibrium (e.g. see **Fig. 3.3B** and **Fig. 3.12**). Using fitted parameters, the fractions of self-assembled conjugates with a certain number of RCs per QD could be assigned, producing a profile that closely matched a Poisson distribution (**Fig. 3.11C**). The model also enabled effects stemming from variation of sample concentration to be accounted for. In addition, the model provided a kinetic understanding of why protein-QD conjugation approximated to Poisson statistics and revealed that the variance between the simulation and Poisson statistics could be the consequence of a dependence of the number proteins per QD on previous species. The knowledge generated on RC-QD conjugation could be readily transferred to other protein/nanomaterial self-

assembling systems and, moreover, opens the possibility of decoding heterogeneity in conjugation systems comprising more than just one protein and one nanoparticle.

FRET distances and efficiencies were estimated through a process that took into account RC:QD heterogeneity (**Fig. 3.14**). The FRET efficiencies of the WT_H RC and all BPhe-replacement mutants increased with an increasing RC:QD ratio, with high ratio RC/QD conjugates showing an apparent ~0.6 nm decrease in FRET distance relative to low ratio conjugates (**Fig. 3.14C**). FRET efficiency was also higher in the high ratio conjugates than in the low ratio conjugates (**Fig. 3.14B**). Intriguingly, a transition was observed in both parameters around a RC:QD of 2.5, with less variation before/after this region. A RC:QD ratio of 2.5 also marked the point at which emission from P-deficient $VL157R_H$ RCs was maximal, and I suggest that these trends are manifestations of the influence of increasing crowdedness in the shell of bound proteins. In the case of $VL157R_H$ RCs, emission increases up to a RC:QD of 2.5 as more energy is passed from the QD to the surrounding RCs. Beyond this point increased protein crowding causes the mean inter-protein distance to fall below the exciton diffusion radius and multiple excited states could co-exist in a single conjugate with multiple RCs, allowing annihilation to occur and causing the amount of emission to progressively drop. One possibility is that the effect of crowding on FRET efficiency and distance could be attributable to the orientation factor, κ^2, that used to describe dipole-dipole coupling between donor and acceptor. For the purposes of calculating E_{DA} and R a value of ⅔ was assigned to κ^2, but it may be that this (commonly applied) simplification becomes less valid when the distribution of possible orientations between the QD and the four acceptor bacteriochlorins in each RC becomes more constrained in progressively more densely-packed RC/QD conjugates at the higher RC:QD ratios. Another speculation depends on the energy processing capability of RC with $P870^+$ or ground state[219] as at low RC:QD ratios, the population of oxidized RC would be higher than at the high RC:QD ratios as the result of energy transfer partition.

3.4. Conclusions

The work described in this chapter established that water-soluble CdTe QDs form a specific binding interaction with *Rba. sphaeroides* RCs when the latter are modified by an extra-membrane poly-histidine tag, the tag also serving to orient RCs at the QD surface. The dissociation constant between a single RC and QD was estimated as being 8.1 nM, indicating a tight binding interaction, and this interaction was well explained by a model assuming multiple, independent binding events. Monodispersed conjugates were directly visualized by negative stain TEM and were found to have a mean diameter of ~ 21.4 ± 1.0 nm. Bound RCs quenched QD fluorescence and, conversely, QD excitation drove photochemistry in WT or

carotenoid-less RCs and drove RC emission in photochemically-inactive RCs lacking the primary electron donor BChls. Quenching of QD emission was sensitive to the spectral overlap with RC absorbance, consistent with a FRET mechanism for the energy transfer. The estimated FRET distance *R* was consistent with morphologies of the RC-QD conjugates predicted from modelling of the RC/QD interaction and measured through experiment. The single donor/acceptor FRET efficiency, E_{DA}, was of the order of 0.53 for the smallest conjugates involving WT_H RCs, and somewhat higher (0.55-0.6) for cofactor-replacement mutant RCs with enhanced spectral overlap. A decrease in *R* and increase in E_{DA} was seen in larger RC/QD conjugates, effects that are attributed to increased packing of RCs around a central QD that constrains their dynamic freedom. Evidence suggestive of exciton-exciton annihilation was also seen when photochemically-inactive RCs were packed around a central QD at a high RC:QD ratio. It is concluded that stable conjugates of a well-defined composition can be formed between His-tagged RCs and water-soluble RCs. In addition to the QDs acting as a synthetic antenna to drive RC photochemistry, they also have potential to act as a hub for the assembly of more complex photosystems involving novel combinations of natural and synthetic components, as explored in Chapter 4.

3.5. Supplementary Figures

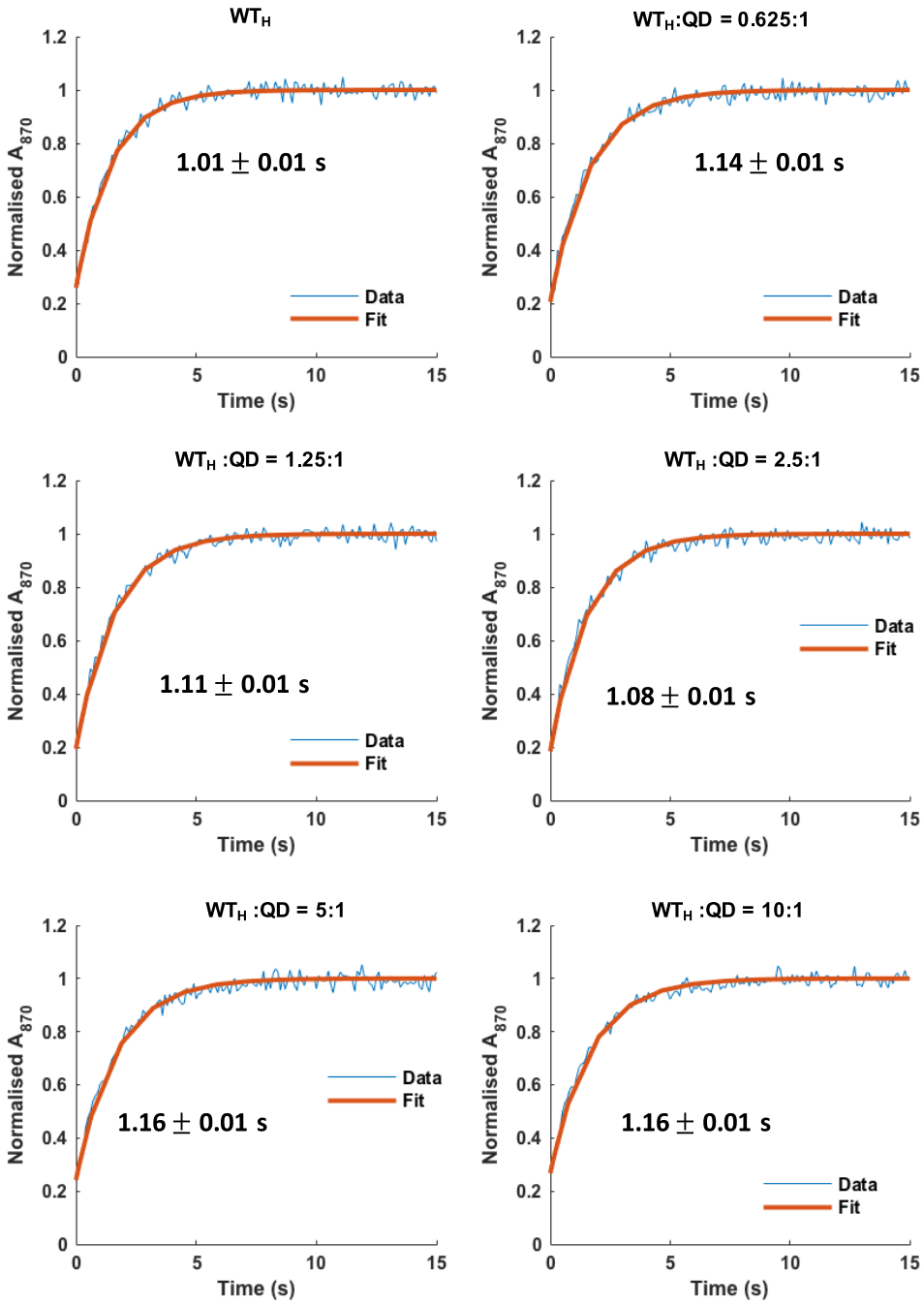


Figure S3.1. Kinetics of dark recovery of strong P photobleaching at 870 nm. Each panel shows averaged kinetic traces (n = 5) and a fit using **Eq. 2.1**. Materials used were QDs and WT_{His} RCs. The composition of each sample is indicated above each panel. [RC] was 5 μM. White light excitation was applied for 0.5 seconds.

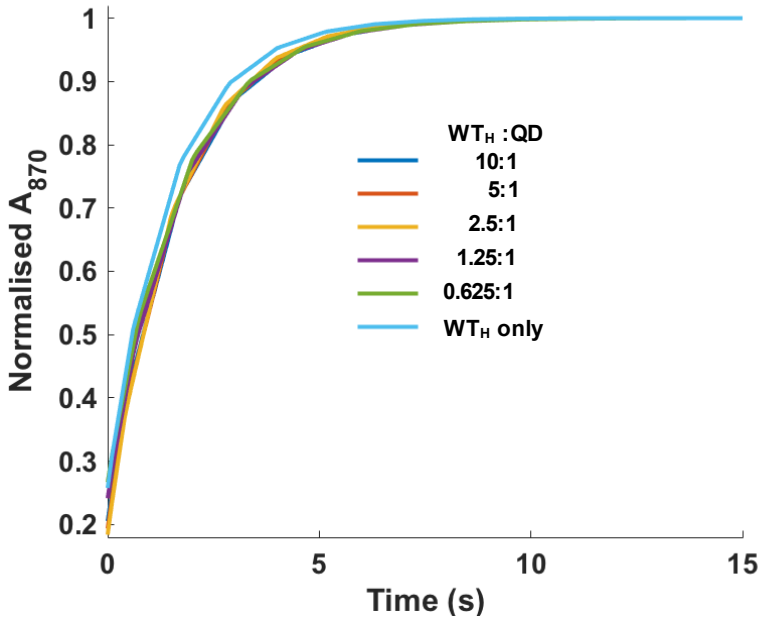


Figure S3.2. Binding to QDs does not affect the kinetics of RC $P^+Q_B^-$ charge recombination. Overlay of fitted kinetics of dark recovery of strong P photobleaching at 870 nm following 0.5 s of white light excitation of WT_{His} RCs and QDs at the indicated ratios.

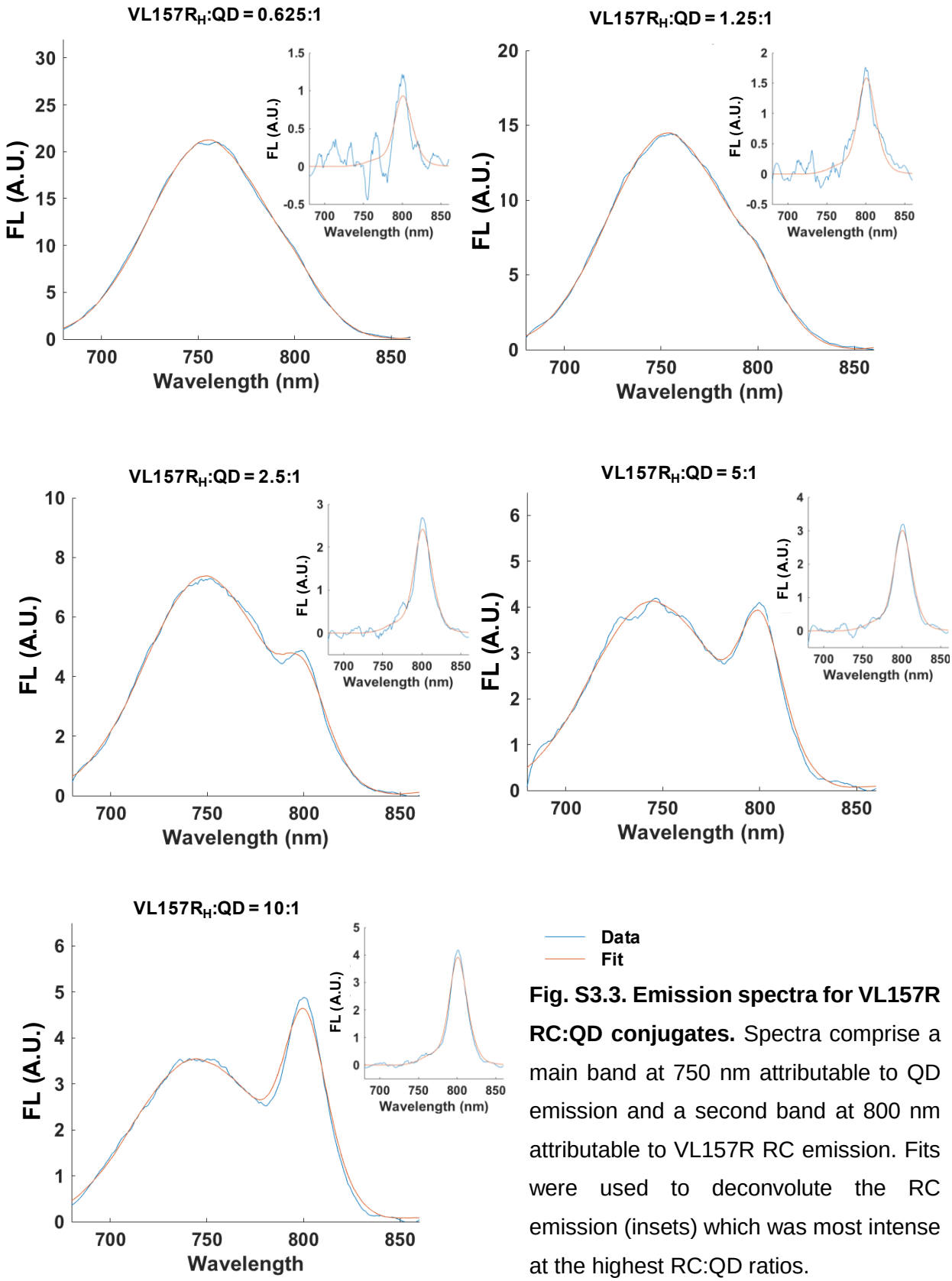


Fig. S3.3. Emission spectra for VL157R RC:QD conjugates. Spectra comprise a main band at 750 nm attributable to QD emission and a second band at 800 nm attributable to VL157R RC emission. Fits were used to deconvolute the RC emission (insets) which was most intense at the highest RC:QD ratios.

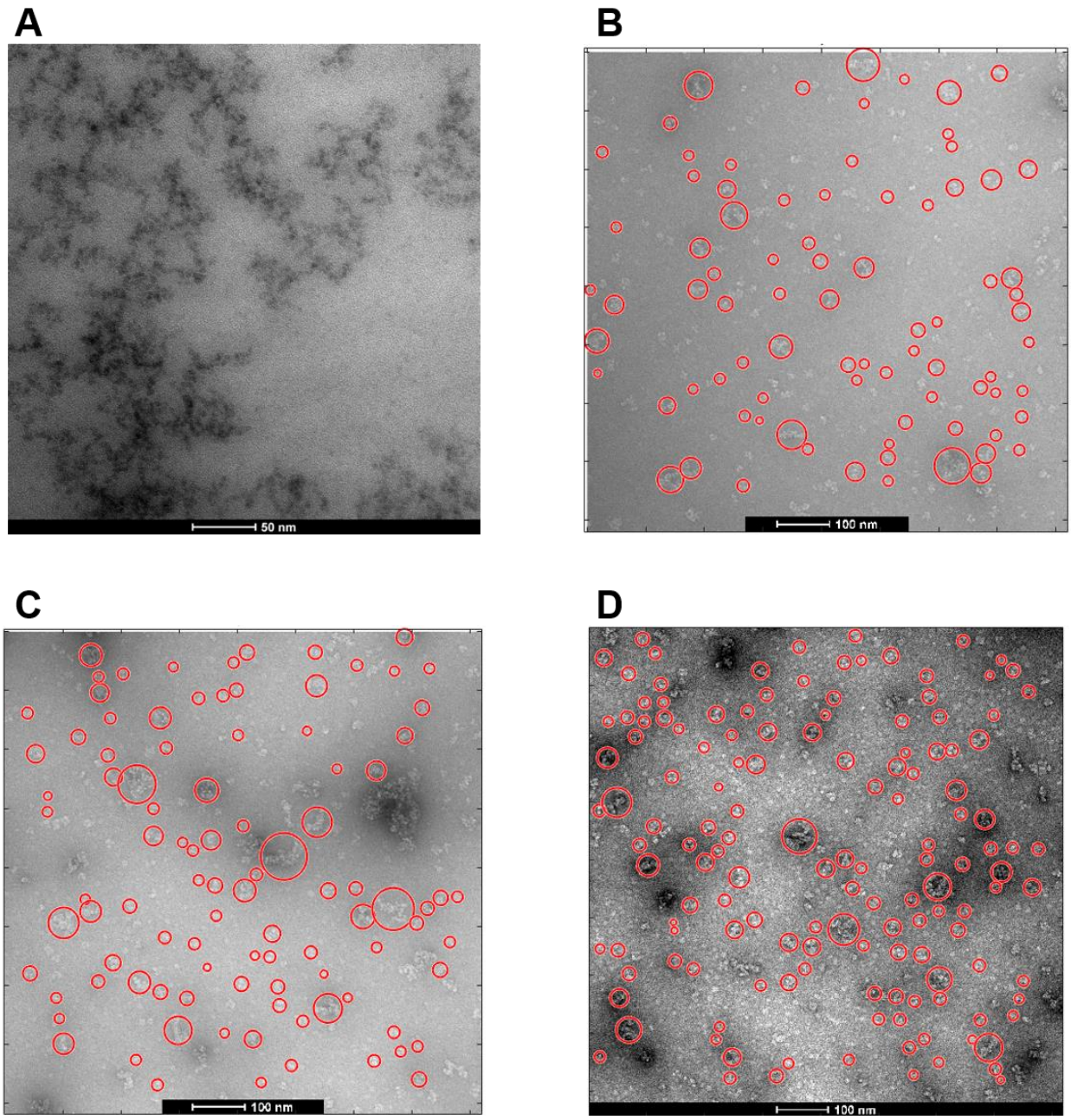


Figure S3.4. TEM images. (A) QD-only sample. (B-D) Negative stained conjugates formed from a 10:1 mix of WT_{His} RCs (1 μM) and QDs (100 nM).

Chapter 4

Minding the Gap between Plant and Bacterial Photosynthesis for Solar Energy Conversion

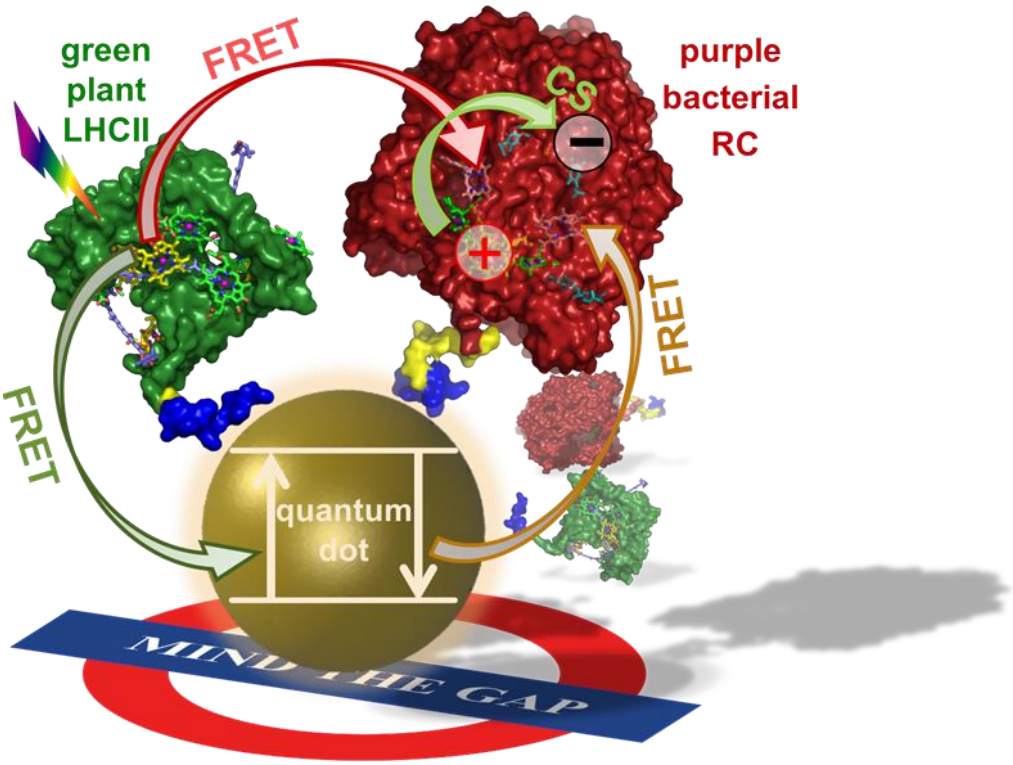
A manuscript describing the results presented in the chapter has been submitted for publication and was written by me:

Liu, J., Mantell, J. and Jones, M.R. (2019) Minding the gap between plant and bacterial photosynthesis within self-assembling biohybrid nanoconjugates for solar energy conversion.

The experiments and analysis described in this chapter were performed by me. Mrs. J. Mantell assisted with the collection of transmission electron microscopy images. XylE protein was a gift from Dr. Natalie D. Bartolo.

Abstract

Many current strategies for meeting mankind's future energy demands through the exploitation of plentiful solar energy have been influenced by the efficient and sustainable conversions in natural photosynthesis. A limitation affecting the solar energy conversion based on photosynthetic proteins is the selective spectral coverage that is the consequence of their particular natural pigmentation. We demonstrate a semi-synthetic polychromatic photosystem that displays an overall efficiency comparable to that seen in natural photosystems by bottom-up self-assembling of chlorophyll-based major light harvesting complex (LHCII) from the oxygenic green plant *Arabidopsis thaliana*, the bacteriochlorophyll-based photochemical reaction centre (RC) from the anoxygenic purple bacterium *Rhodobacter sphaeroides* and artificial material quantum dots (QDs). Polyhistidine tag adaptation of LHCII and RC enables self-assembly of LHCII/RC/QD nanoconjugates, thermodynamic of which can be well modelled and paramterised. The tricomponent biohybrid nanoconjugates displayed enhanced solar energy conversion via either direct energy transfer or indirect pathway offered by QD.



4.1. Introduction

As the largest sustainable source of energy, use of solar radiation could address mankind's growing energy needs through sustainable and environmentally-benign approaches[127]. Providing inspiration for how to achieve this, natural photosynthesis converts solar energy at a rate that is several times greater than our current energy demand[91]. A feature of natural photosynthesis is that different organisms exploit discrete parts of the available solar spectrum depending on their pigmentation. Plants, algae and cyanobacteria contain Chl as their principal photosynthetic pigment (**Fig. 4.1A,** left) and exploit much of the visible spectrum between 400 and 700 nm, which accounts for around 50% of the solar energy arriving at the earth's surface[125]. Complementary to this, diverse anoxygenic photosynthetic bacteria contain BChl as their primary photosynthetic pigment (**Fig. 4.1A**, right) and utilize sunlight most strongly in the near-UV and near-infrared spectral regions[34]. Overcoming the limitations of natural pigment palettes and enabling expanded solar energy conversion are at the heart of photosynthesis research themes [127,220,221] and are important considerations for many biohybrid photovoltaic or photoelectrochemical chemical devices[105,107,109,115,120,222,223].

In this chapter, a self-assembling tri-component biohybrid photosystem is described that combines *Rba. sphaeroides* RCs (**Fig. 4.1B**), CdTe QDs and the major LHCII from *A. thaliana* (**Fig. 4.1C**), three components that display complementary spectral coverage (**Fig. 4.2A**). The two proteins have no natural propensity to associate or work together for solar energy conversion, but when adapted with His-tags they can be induced to self-assemble on the surface of a QD into a tuneable LHCII/RC/QD conjugate. We show that, as with the RC/QD conjugates described in Chapter 3, the interface between LHCII and the QD is solely provided by the His-tag and the thermodynamics of conjugation can be well-described by a multiple/independent binding model between LHCIIs, RCs and the QD hub[224]. This enables predictable self-assembly of conjugates with defined and tuneable compositions. By selecting QDs with suitable properties (6.5 nm diameter, 750 nm emission) they not only provided a physical hub for photosystem assembly but also bridged the energy gap between LHCII and the RC through the spectral overlap highlighted by the shaded areas in **Fig. 4.2A,B**[155]. The design of the system (**Fig. 4.2B**) envisaged that excitation of LHCII initiated the delivery of energy from the Chl light harvesting system of LHCII to the BChl charge separation system of the RC, either directly or via the energy conduit provided by the QD. The observed energy flows in this tri-component system reflect strategies within natural photosystems that ensure high efficiency and plasticity to changes in environmental conditions such as light quality[220,221].

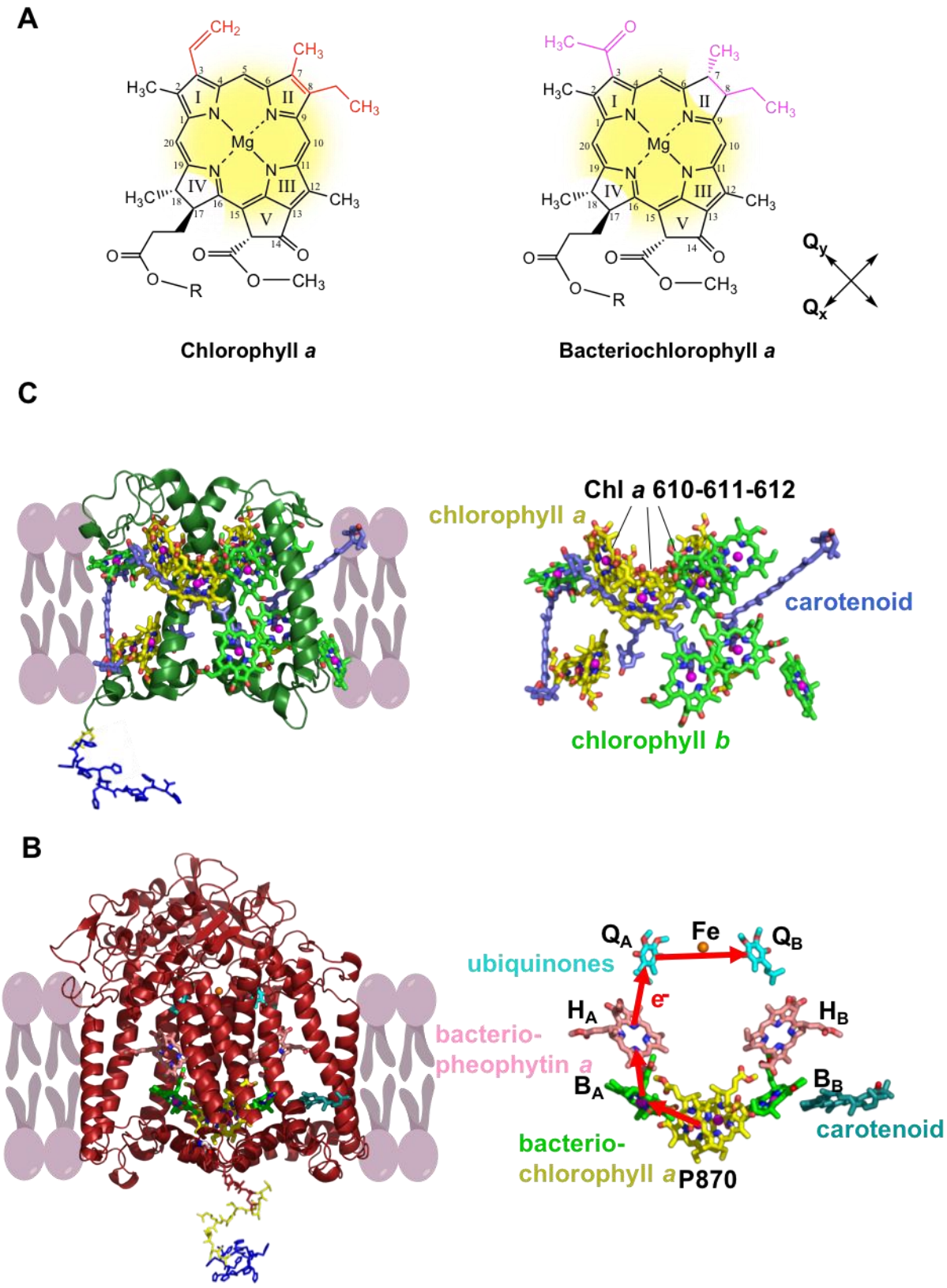


Figure 4.1. Pigment-protein structure and the mechanism of solar energy conversion. (A) Differences affecting π-conjugation in Chl *a* and BChl *a* are highlighted in red and purple, respectively). The directions of the Q_y and Q_x transition dipoles are indicated by arrows that

cross at the central magnesium. R represents the hydrocarbon side chain. Differences in bonding affect the shape and symmetry of the conjugated π-electron system (yellow shading). (B) The cofactors of the *Rba. sphaeroides* RC comprise four BChl *a* (yellow or green carbons), two BPhe *a* (pink carbons), two ubiquinone-10 (cyan carbons), one carotenoid (teal carbons) and one Fe atom (orange sphere). Four-step charge separation (red arrows) takes place between two P870 BChls (yellow carbons) and the dissociable Q_B ubiquinone. Other atom representations are: nitrogen – blue; oxygen – red; iron - brown sphere, magnesium – magenta sphere. The cofactors are held in place by a largely intra-membrane protein scaffold (maroon ribbons). The His-tags and linkers are shown as blue and yellow sticks, respectively. (C) The cofactors of LHCII comprise eight Chl *a* (yellow carbons), six Chl *b* (green carbons) and four carotenoids (slate carbons). The cofactors are held in place by a three-helix intra-membrane scaffold. The His-tags and linkers are shown as blue and yellow sticks, respectively. The Chl *a* 610-611-612 cluster forms the lowest energy state within LHCII monomer and energy harvested by other pigments quickly relaxes to this emitting site.

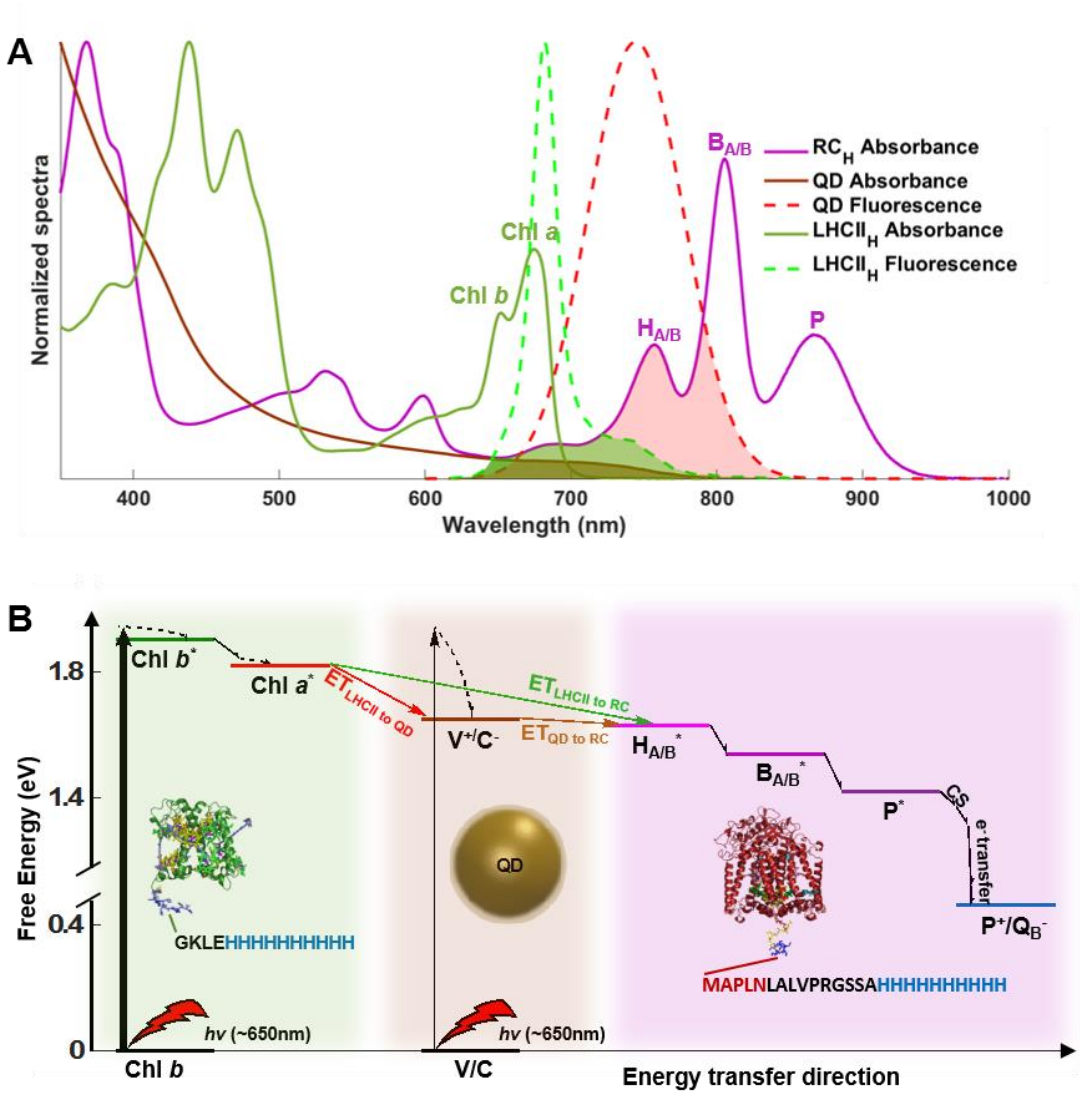


Figure 4.2. Tri-component conjugate design. (A) Normalized absorbance and emission spectra for conjugate components. Spectral overlap between $LHCII_H$ fluorescence and RC_H absorbance (9.34×10^{-13} cm^{-3}) is shown in light green shadow, that between $LHCII_H$ fluorescence and QD absorbance (3.77×10^{-12} cm^{-3}) in green/brown shadow, and that between QD fluorescence and RC_H absorbance (3.78×10^{-12} cm^{-3}) in red/green shadow. (B) Plot of the envisaged energy transfer (ET) scheme in a tri-component conjugate. Light at 650 nm should excite mostly Chl *b* in $LHCII_H$, with subsequent exciton transfer to the lowest Chl *a* energy state[67,225]. Further energy transfer to the RC can take either a direct pathway (green arrow) or indirect pathway via the QD (red/brown arrows), triggering RC internal energy transfer followed by charge separation (CS). In $LHCII_H$ the His_{10}-tag (blue letters) is attached to the protein's C-terminus (green letters) via a linker (black letters). In RC_H a linker connects the His_{10}-tag to the C-terminus of PufM (maroon letters).

4.2. Results

4.2.1. Assembly and characterisation of di-component LHCII/QD conjugates

The first step in the development of tri-component LHCII/RC/QD conjugates was to assemble and characterise di-component conjugates formed between LHCII and CdTe QDs. To enable genetic modification of *A. thaliana* LHCII a gene encoding the LHCII apoprotein was expressed in *E. coli* and the holoprotein refolded with pigments *in vitro*.

The starting point for the production of recombinant LHCII variants was a pET-28a vector containing a gene encoding the Lhcb1.3 protein from *A. thaliana* (UniProtKB entry P04778), that was a kind gift from Prof. Roberta Croce, Vrije Universiteit Amsterdam. Modification of this gene was carried out by Gibson assembly using oligonucleotides sourced from Eurofins or using the Q5® Site-Directed Mutagenesis Kit from NEB, as described in Section 2.4. Designed apoproteins were expressed in *E. coli* Rosetta™ 2 (Novagen), purified from inclusion bodies, and the mature pigment-protein assembled in vitro using purified Chl *a*, Chl *b* and carotenoid pigments with DDM as the supporting detergent, as described in Section 2.6

A version of this *A. thaliana* LHCII gene encoding a Lhcb1.3 polypeptide lacking twelve amino acids from the N-terminus (denoted dLhcb1) was further modified with a deca-histidine tag at its C-terminus. These twelve amino acids are not resolved in available X-ray crystal structures of LHCII[64] (1RWT), and their removal does not affect the functional properties of isolated LHCII[151]. The resulting polypeptide was denoted dLhcb1-H_{10} **(Fig. 4.3A)**. This was expressed in *E. coli* and the mature holoprotein, named $LHCII_H$, refolded. For comparison, native LHCII was purified from spinach (see procedures in Section 2.6.4)

Analysis by ultracentrifugation of 0%/25%/60% sucrose density gradients (Section 2.14) showed that this His_{10}-tag modification produced approximately 90 % binding to QDs, whereas almost no binding was seen for native LHCII purified from spinach (**Fig. 4.3B**). This paralleled strong and specific binding of His_{10}-tag-modified RCs (for convenience named RC_H in this chapter, where only proteins with WT pigment compositions are used) to the same QDs as shown in Chapter 3 (**Fig. 4.3B**). When mixed in a 1:1 ratio, both $LHCII_H$ and RC_H bound strongly to QDs at the same total-protein to QD ratio (**Fig. 4.3B**).

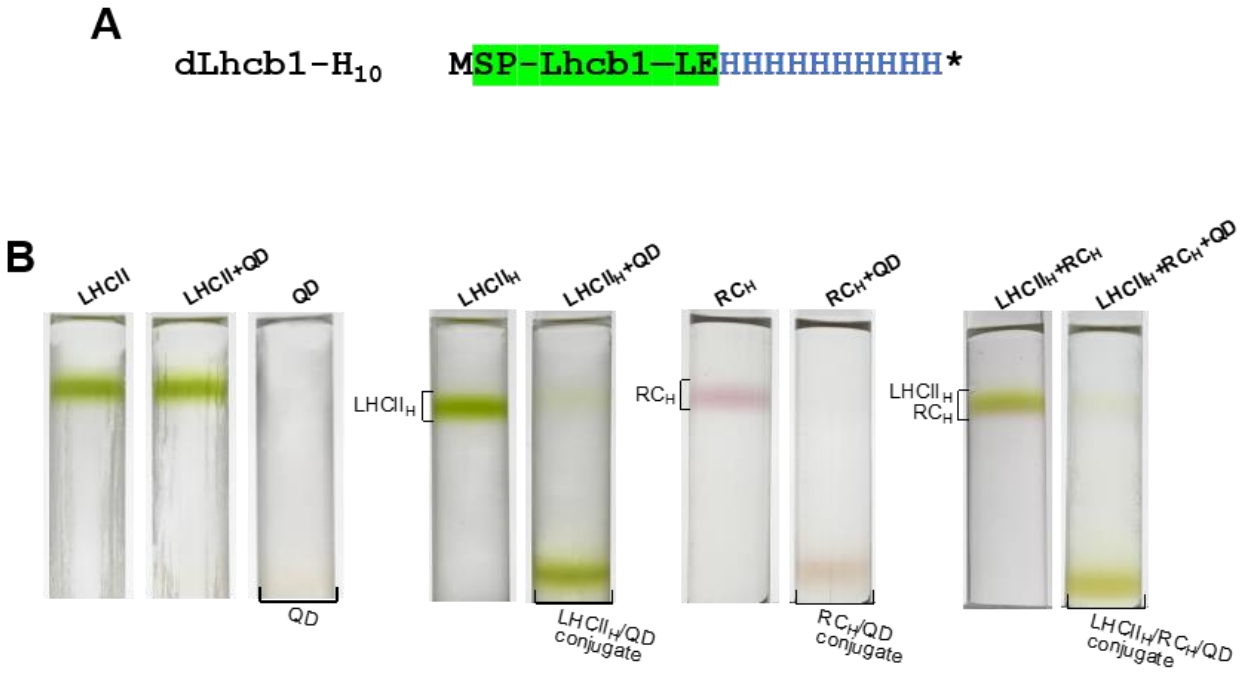


Figure 4.3. Binding of LHCII$_H$ to QDs. (A) Design of the dLhcb1-H$_{10}$ apoprotein from which the LHCIIH holoprotein is refolded. The N-terminal methionine is followed by a truncated Lhcb1 polypeptide (first two and last two residues shown) and then a His$_{10}$ tag. (B) Assays of QD binding by sucrose density gradient ultracentrifugation of (left to right) spinach LHCII, LHCII$_H$, RC$_H$ and a 1:1 LHCII$_H$:RC$_H$ mix in the absence and presence of QDs (10:1 molar ratio of total-protein:QD). In the absence of binding, proteins (green or red band) migrate to the upper 0%/25 % interface whereas QDs (faint brown band) migrate to the lower 25%/60 % interface. Proteins are pulled down to the lower interface when bound to a QD. At this 10:1 molar ratio almost all protein molecules modified with a His-tag bind to a QD and migrate to the lower interface. The unbound fraction mainly comprised damaged proteins unable to bind.

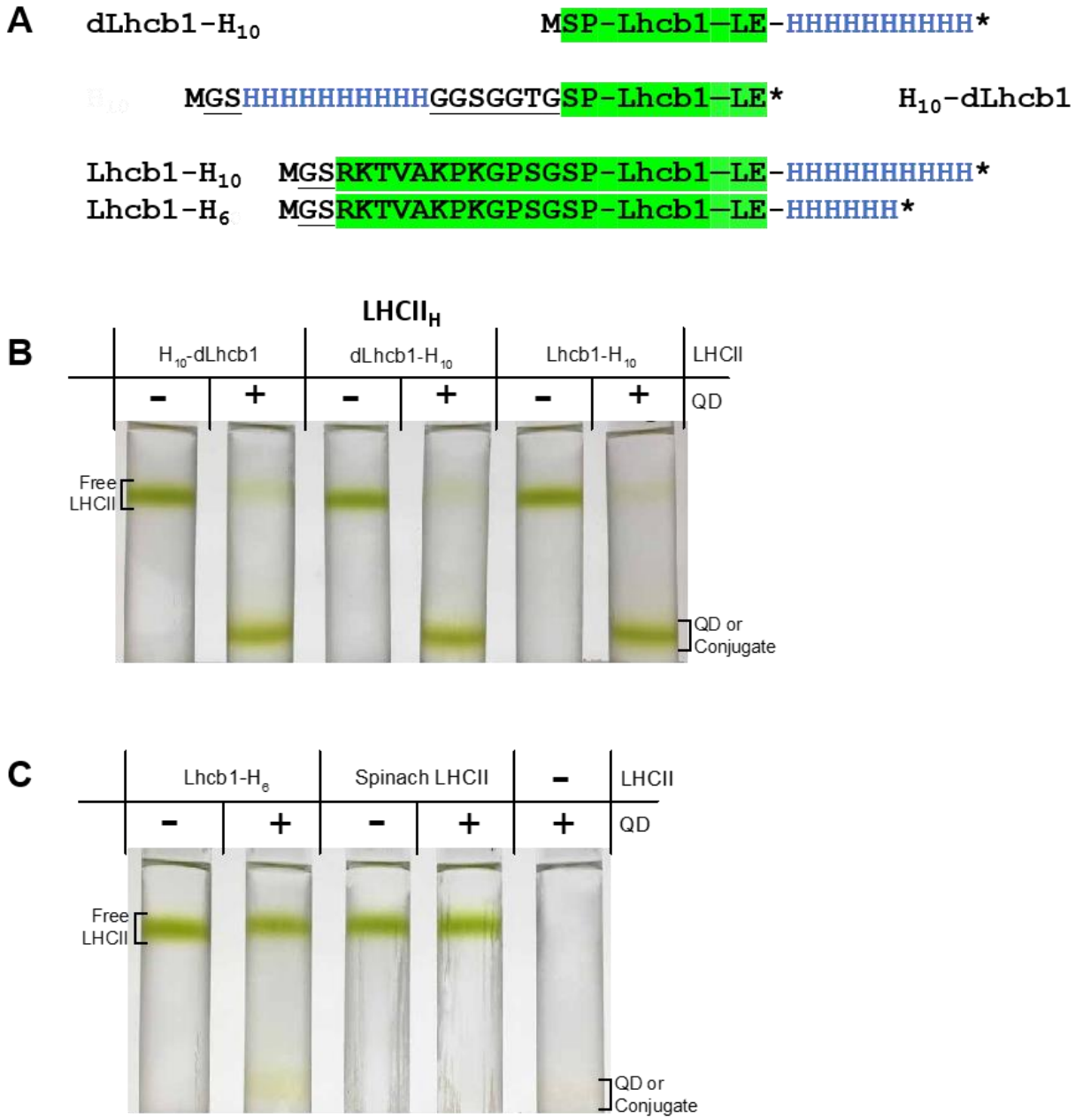


Figure 4.4. Effect of His-tag position/length on binding of LHCII to QDs. (A) Designs of all four LHCII constructs. The Lhcb1 polypeptide is highlighted in green and linkers are underlined. (B) Assays of QD binding by sucrose density gradient ultracentrifugation. Strong (~90 %) binding is conferred by a His_{10}-tag. Binding was not affected by removal of the 12 N-terminal residues of Lhcb1 or the position of the His_{10}-tag at the N- or C-terminus. (C) Very little binding was seen between native spinach LHCII complexes and QDs (~5%). Moderate binding is conferred by a His_6-tag (~30 %). In (B), the complex also labelled $LHCII_H$ is the same as that shown in **Fig. 4.3B**.

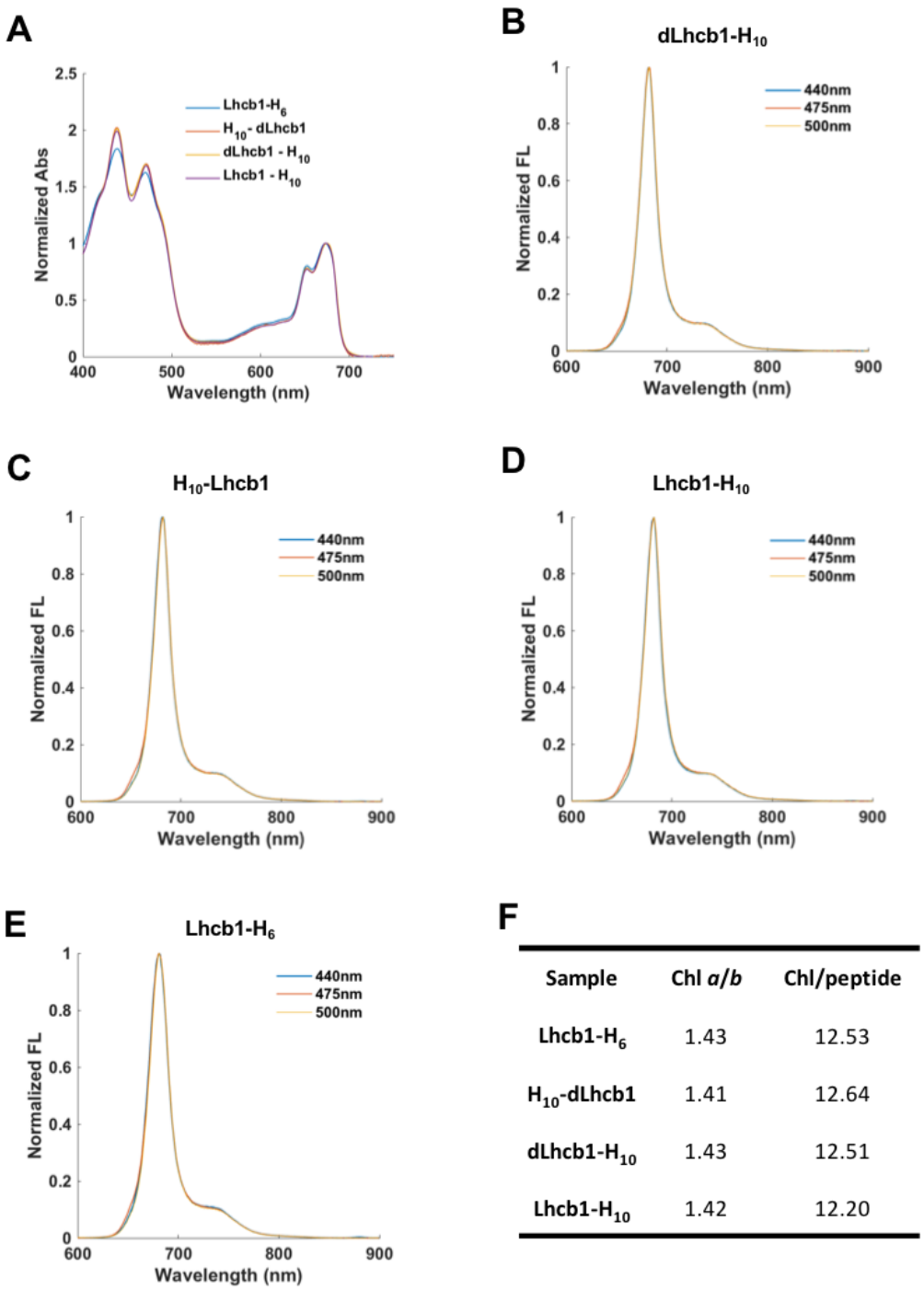
To look at any dependence on the position or length of the His-tag, a further three modified LHCII complexes were engineered also based on the Lhcb1.3 polypeptide from *A. thaliana* (**Fig. 4.4A**). Two had either a six-His or ten-His tag added at the C-terminus of the full Lhcb1 polypeptide (termed Lhcb1-H_6 and Lhcb1-H_{10}, respectively). The third was based on the version of Lhcb1.3 with twelve amino acids removed from the N-terminus and had a ten-His tag added to this truncated N-terminus via a linker (H_{10}-dLhcb1).

Each of these three additional modified Lhcb1 apoproteins was also expressed in *E. coli* and the mature holoprotein refolded. Analysis of all four recombinant LHCII by sucrose density gradient ultracentrifugation showed that modification of either the full-length or N-terminally truncated LHCII with a His_{10}-tag produced strong binding to CdTe QDs (~90 % - **Fig. 4.4B**), whereas modification with a His_6-tag produced markedly weaker binding (~30 % - **Fig. 4.4C**) and native LHCII hardly associated with QDs (~5 % - **Fig. 4.4C**). Estimates of percentage binding were based on absorbance spectra of fractions from deconstructed sucrose density gradients. It has previously been reported[151] that the N-terminal region of LHCII contains several basic amino acids that can significantly contribute to binding to water-soluble QDs with an anionic capping layer (mercaptopropionic acid was used in the present study to provide an anionic capping layer), but our observations were that removal of this region in dLhcb1-H_{10} and H_{10}-dLhcb1 had no effect on binding relative to that seen for Lhcb1-H_{10} (**Fig. 4.4C**) and the strength of binding was also not dependent on placement of the His_{10}-tag at the N- or C-terminus (**Fig. 4.4C**). It was, however, clearly sensitive to the length of the His-tag.

1.1.1. Quality of refolded LHCII

After refolding, all four recombinant LHCIIs exhibited similar absorption profiles (**Fig. 4.5A**). All emission spectra had a maximum around 682 nm and a line shape that was independent of excitation wavelength (**Fig. 4.5B-E**), indicating the assembly of correctly folded complexes. Chl *a*/Chl *b* ratios for the reconstituted monomeric LHCIIs (**Fig. 4.5F**) were consistent with previously published values while total Chl to peptide molar ratios (**Fig. 4.5F**) were similar to one another and again agreed with literature values [145,179,226]. Mean values of 1.42 for Chl *a*:Chl *b* and 12.5 for total Chls per complex implied that every reconstituted LHCII contained around 7.3 Chl *a* and 5.2 Chl *b*, in close agreement with estimations from native purified complexes[227].

The quantum yield of refolded $LHCII_H$ was determined by comparison with the dye DCM, as described in Section 2.11.2 and summarised in **Fig. 4.6.**



F

Sample	Chl *a*/*b*	Chl/peptide
Lhcb1-H_6	1.43	12.53
H_{10}-dLhcb1	1.41	12.64
dLhcb1-H_{10}	1.43	12.51
Lhcb1-H_{10}	1.42	12.20

Figure 4.5. Quality control of Refolded LHCII. (A) Absorbance spectra of the four refolded LHCII. (B-E) Emission spectra of the four refolded LHCII. In all cases the line shape of the emission spectrum was independent of excitation wavelength, which is diagnostic of a correctly-folded complex. (F) Chl compositions of the four refolded LHCII based on spectra of 80 % acetone extracts and a protein extinction coefficient of 546,000 $M^{-1}cm^{-1}$.

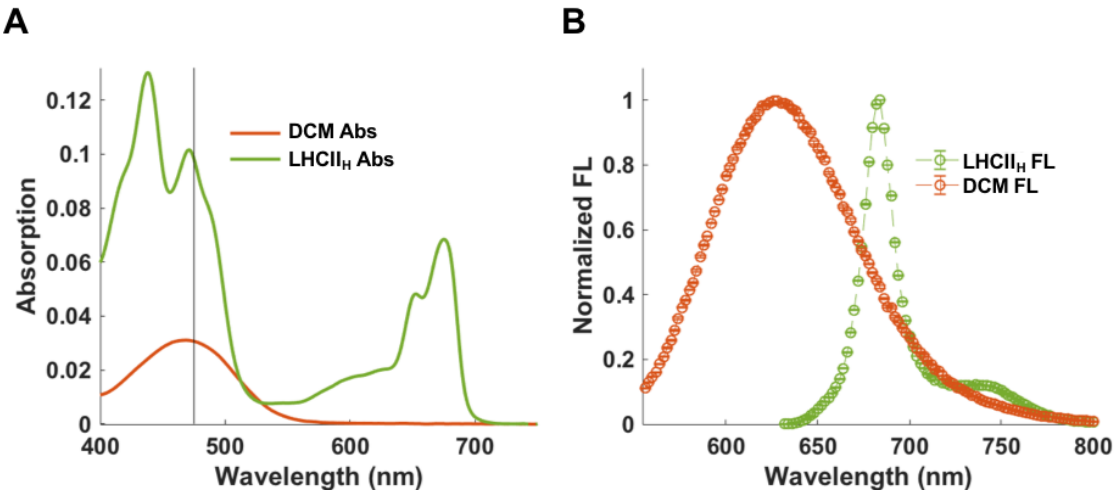


Figure 4.6. Determination of the quantum yield of LHCII$_H$. (A) Absorption spectra of 100 nM solutions of LHCII$_H$ and DCM. Excitation was at 475 nm (black line). (B) Normalized emission spectra of the same samples after calibration of the emission response[185].

4.2.2. Assembly of $LHCII_H$/QD conjugates and evidence for $LHCII_H$ to QD energy transfer

As with the RC_H/QD conjugates described in Chapter 3, experimental data on binding of $LHCII_H$ to QDs from spectroscopic analysis of fractionated sucrose density gradients (**Fig. 4.3B**) could be well described by a model employing multiple independent binding sites[200] that is summarised in **Fig. 4.7A**. Details of this modelling are given in supplementary Section 4.6.1 at the end of this chapter. This model was the same as that used for RCH/QD conjugates, involving the application of **Eqs. 3.1-3.7** described in Chapter 3.

The best fit of the model to the data (**Fig. 4.7B**) produced a micro-kinetic constant (k_{micro}) of 7.19 μM^{-1} and a maximum permitted binding number ($\bar{n}$) of 17. A K_d of 8.2 nM for the first $LHCII_H$ to bind to a QD was deduced from this information using **Eq. 3.2** (see Section 4.6.1). This was nearly identical to a value of 8.1 nM obtained in Chapter 3 for the binding of RCs with the same deca-His tag to the same QDs.

Titration of QDs with an increasing molar ratio of $LHCII_H$ produced a decrease in $LHCII_H$ emission relative to concentration-matched $LHCII_H$-only controls (**Fig. 4.7C,** solid green vs dashed green**)** and a corresponding increase in QD emission (**Fig. 4.7C,** solid brown), diagnostic of FRET to the QDs from the attached $LHCII_H$ (**Fig. 4.7D**). Amounts of $LHCII_H$ and QD emission were quantified by the deconvolution of emission spectra (see **Fig. S4.1** for sample data).

4.2.3. Energy transfer efficiency and FRET distance for $LHCII_H$/QD conjugates.

Single donor-acceptor energy transfer efficiencies were estimated by combining thermodynamic data on $LHCII_H$/QD conjugation with data on $LHCII_H$ and QD emission from conjugates with various $LHCII_H$:QD stoichiometries (**Fig. 4.7C**). In these conjugates multiple copies of $LHCII_H$ donated energy to an associated single QD (**Fig. 4.7D**), which could be explained by a multiple donor-single acceptor FRET mechanism.

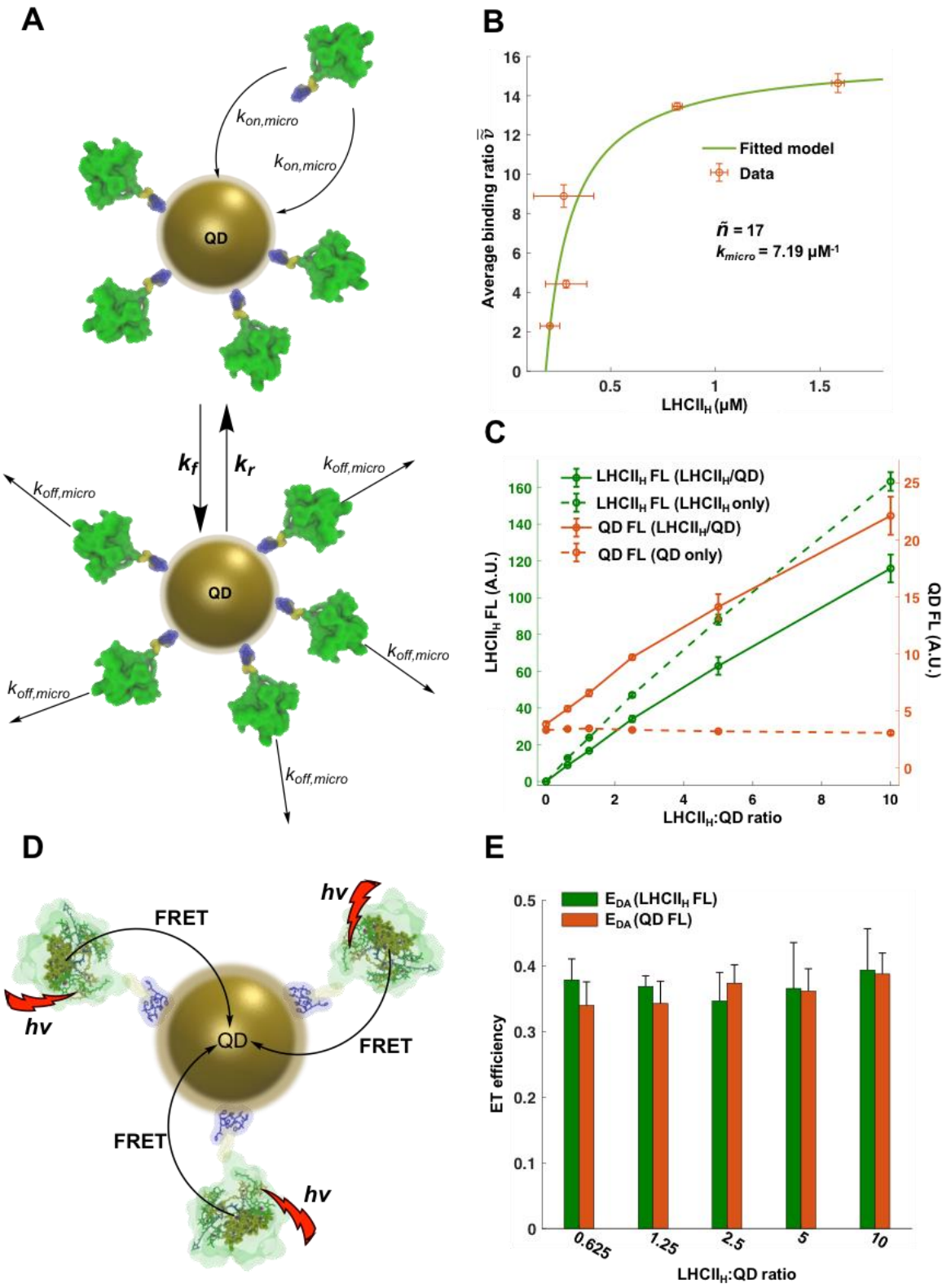


Figure 4.7. Binding and energy transfer in LHCII$_H$/QD conjugates. (A) Scheme for the interaction between multiple LHCII$_H$ and a QD. Microscopic constants $k_{on,micro}$ and $k_{off,micro}$ describe association/dissociation events between a single LHCII$_H$ and the QD. Parameters k_f and k_r indicate the macroscopic kinetic outcomes that are the consequence of all possible microscopic events (see **Eqs. 3.6-3.7** in Section 3.2.7). (B) Experimental data on average

binding ratio as a function of the concentration of unbound $LHCII_H$ were fitted using **Eq. 3.1,** as detailed in section **4.6.1**. (C) Deconvoluted $LHCII_H$ and QD fluorescence as a function of the $LHCII_H$:QD ratio in mixes used to form $LHCII_H$/QD conjugates (solid-lines), or in concentration-matched $LHCII_H$-only or QD-only samples (dashed-lines). Decreased $LHCII_H$ fluorescence and increased QD fluorescence in conjugates indicates $LHCII_H$→QD FRET. (D) FRET scheme between multiple bound $LHCII_H$ and a QD. Clusters of three low energy Chl *a*, likely the FRET donor, are highlighted in spheres. Due to the QD's ability to accommodate multiple $LHCII_H$ harvesters the absorbance cross-section of the conjugate was greatly enhanced, boosting QD emission in response to 650 nm excitation to nearly six-fold that recorded for naked QDs at a $LHCII_H$:QD equal to 10:1. (E) Estimates of the efficiency of single donor-single acceptor ET in conjugates formed from mixes with different $LHCII_H$:QD ratios.

The single donor to acceptor ET efficiency, E_{DA}, was estimated independently from either $LHCII_H$ emission ($E_{DA(LHCII)}$) or QD emission ($E_{DA(QD)}$) using:

$$E_{DA(LHCII)} = \frac{[LHCII]_T}{[LHCII]_B}\left(1 - \frac{F_{Conj(LHCII)}}{F_{LHCII}}\right) \quad \textbf{(4.1)}$$

$$E_{DA(QD)} = \frac{F_{Conj(QD)} - F_{QD}}{F_{QD}} \frac{1 - 10^{-Abs_{QD}}}{1 - 10^{-\frac{[LHCII]_B}{[LHCII]_T} Abs_{LHCII}}}} \quad \textbf{(4.2)}$$

where $F_{Conj(LHCII)}$ and $F_{Conj(QD)}$ represented $LHCII_H$ and QD fluorescence, respectively, in conjugates and F_{LHCII} and F_{QD} represented the fluorescence of concentration-matched $LHCII_H$-only and QD-only controls. The fraction of the total concentration of $LHCII_H$ ($[LHCII]_T$) that was bound to the QDs ($[LHCII]_B$) was deduced from the final steady state of the simulation described above (and in detail in Section 4.6.1). Abs_{LHCII} and Abs_{QD} were the absorbances of $LHCII_H$ and QDs, respectively, at the 651 nm excitation wavelength, measured after quenching experiment.

Eqs. 4.1 and **4.2** were used to calculate values of $E_{DA(LHCII)}$ and $E_{DA(QD)}$ for $LHCII_H$/QD conjugates formed in five mixes with $LHCII_H$:QD molar ratios varying between 0.625 and 10. A mean value for $E_{DA(LHCII)}$ of 37.1 ± 1.7 % agreed well with a mean value of 36.1 ± 2.0 % for $E_{DA(QD)}$, individual values showing no significant dependence on the $LHCII_H$:QD ratio (**Fig. 4.7E**). This agreement between independent estimates reinforced the conclusion that the mechanism of energy transfer from $LHCII_H$ to the QD was FRET.

The distance for 50 % FRET efficiency, R_0, was determined from **Eq. 3.9** (see Chapter 3), where Φ_D was the $LHCII_H$ quantum yield, J was the spectral overlap between $LHCII_H$ fluorescence and QD absorbance, and n was the refractive index of water (1.33). The spectral overlap was quantified using the spectra shown in **Fig. S4.2** which produced a value of 3.77 x 10^{-12} cm^{-3} M^{-1}. The orientation factor κ^2 was assumed to be ⅔.

Having determined E_{DA} and R_0, the actual FRET distance, R, was then calculated using **Eq 3.10**. Values arrived at were 7.37 ± 0.09 nm based on the average $E_{DA(LHCII)}$ and 7.42 ± 0.11 nm based on the average $E_{DA(QD)}$. It was a possibility that the commonly-used value for κ^2 of ⅔ may not have been applicable in this system since anchorage of $LHCII_H$ to the QD constrained their relative motion, and hence the two "dipoles" were not freely sampling all orientations, which is the assumption behind κ^2 being assigned a value of ⅔. Nevertheless, this value was used in the absence of an obvious alternative.

These independent estimates of actual FRET distance gave $LHCII_H$/QD conjugate diameters of 19.74 nm or 19.84 nm, with a mean value for R of 7.4 ± 0.1 nm, giving a diameter of 19.8 ± 0.2 nm. These were based on the assumption that the FRET distance is from the centre of the 6.5 nm diameter QD to the centre of $LHCII_H$. An alternative possibility is that the centre of the $LHCII_H$ dipole is not the geometric centre of protein but rather the centre of the emitting red-form Chl *a* 610-611-612 cluster[65] which is located closer to N-terminal side of the protein. As the His_{10}-tag is on the C-terminus of LHCII this moves the FRET donor further from the QD surface and would reduce the estimates of conjugate diameter somewhat, bringing them closer to the value of 19.3 ± 1.0 nm estimated from TEM and image analysis described below.

4.2.4. Architecture of $LHCII_H$/QD conjugates

The morphology of the LHCIIH/QD conjugates was examined by TEM, as described in Section 2.12. This revealed that the addition of QDs to His-tagged protein (1:10 molar ratio) produced a marked change in sample morphology from monodispersed small (<10 nm) objects to a smaller number of larger (>15 nm) objects, consistent with conjugate formation (raw images shown in **Fig. S4.3A**). **Fig. 4.8A** shows a typical multi-lobed object obtained on mixing $LHCII_H$ with QDs, whilst **Fig. 4.8B** shows a schematic of a $LHCII_H$/QD conjugate. Analysis of multiple images (**Fig. S4.3B,C**), carried out as described in Section 2.12.2 produced estimates of conjugate diameter that could be fitted with a lognormal distribution

with a mode at 19.3 ± 1.0 nm (**Fig. 4.8C**) and a variance of 17.5 nm^2, matching well the diameter of 19.8 ± 0.2 nm calculated from FRET analysis.

4.2.5. Architectures of tri-component conjugates

Tri-component conjugates were formed by simply combining RC_H and $LHCII_H$ protein in buffer solution (Tris/DDM) with QDs in various ratios of $LHCII_H$:RC_H and various ratios of total-protein:QD (i.e. $LHCII_H$ + RC_H).

TEM imaging of tri-component conjugates (**Fig. 4.8E**) formed in an equimolar mixture of $LHCII_H$ and RC_H with a total-protein:QD ratio of 10:1 also had a roughly circular profile (**Fig. 4.8D**) and analysis (**Fig. S4.3D,E**) produced a higher mode value of 21.5 ± 1.0 nm (**Fig. 4.8F**) than obtained for $LHCII_H$/QD conjugates, and with a broader distribution (variance of 20.7 nm^2). These were consistent with RC_H being larger than $LHCII_H$. The $LHCII_H$/RC_H/QD diameter mode was very close to a value of 21.4 ± 1.0 nm from an equivalent analysis of RC_H/QD conjugates described in Chapter 3 (section **3.2.5)**. A similar profile was also obtained from TEM image analysis (**Fig. S4.3FG**) of tri-component conjugates comprising $LHCII_H$/$XylE_H$/QD (**Fig. 4.8H**), with a mode of 21.1 ± 1.0 nm and variance of 19.6 nm^2 (**Fig. 4.8I**). $XylE_H$ is a photo-inactive "dummy" protein with the same His_{10}-tag and a similar size to RC_H. The difference in size between the $LHCII_H$/RC_H/QD and $LHCII_H$/$XylE_H$/QD tri-component conjugates and the $LHCII_H$/QD di-component conjugates suggested that although the characteristics of the imaged tri-component conjugates were dominated by the larger protein, the presence of $LHCII_H$ increased the subpopulation of larger objects, presumably due to details of the spatial arrangement in the tri-component conjugate compared to di-component conjugates.

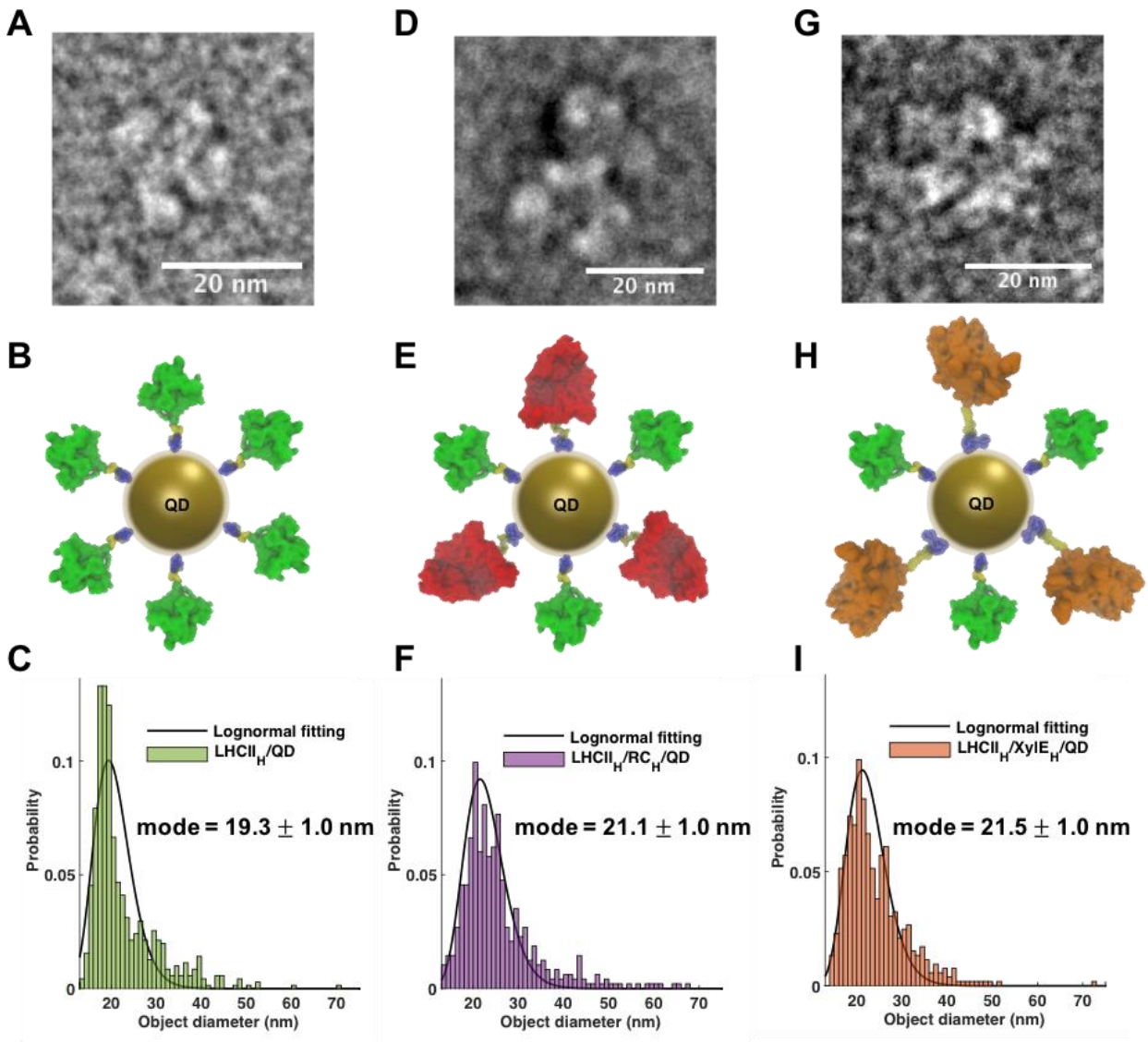


Figure 4.8. Architectures of protein/QD conjugates examined by TEM. (A) An individual $LHCII_H$/QD conjugate from a 10:1 molar mix imaged by TEM. (B) Model of a $LHCII_H$/QD conjugate with $LHCII_H$ in green, C-terminal residues in yellow and the His-tag in blue. (C) Histogram of diameters of $LHCII_H$/QD conjugates determined from image analysis (707 objects), and a fitted lognormal distribution. (D-F) Equivalent data for $LHCII_H$/RC_H/QD tri-component conjugates from a 5:5:1 molar mix (506 objects analysed); RC_H shown in maroon. (G-I) Equivalent data for $LHCII_H$/$XylE_H$/QD tri-component conjugates from a 5:5:1 molar mix (599 objects analysed); $XylE_H$ shown in orange. Flexible linkers are shown in yellow and the His-tags in blue. The model of $XylE_H$ was generated from X-ray crystal structure 4GBZ[228].

4.2.6. Chl to BChl energy transfer in tri-component conjugates.

As it is a quencher, energy transfer to the RC_H from the $LHCII_H$ in tri-component conjugates was detected from bleaching of the 870 nm absorbance band of the primary electron donor BChls that occurs through charge separation. Samples were illuminated for 7 s and the absorbance monitored at 870 nm (see Section 1.17). Kinetic traces were fitted with a simple interconversion model (**Eq. 2.1** in Chapter 2) and a rate of subsequent dark recovery. Measured traces and fits are shown in **Fig. S4.4**, and rates of bleaching and dark recovery are summarised in **Fig. S4.5B,C**. Excitation was through a ~50 nm wide bandpass filter centred at 650 nm, a window where RC_H absorbance is very low compared to that of QDs and, in particular, $LHCII_H$ (**Fig S4.5A**).

Fitted kinetics from these measurements are summarised in **Fig. 4.9A,B**. P870 photooxidation was only modestly enhanced by adding $LHCII_H$ in a 1:1 ratio to either RC_H or otherwise equivalent WT RCs lacking a His-tag (denoted as RC), both in the absence of QDs, or to a mixture of QDs and non-His_{10}-tagged RC (**Fig. 4.9A**). This was the expected result for two proteins in solution that have no natural affinity for one another, and for a mixture with QDs where $LHCII_H$ was able to bind but RC was not due to the lack of a His-tag. In contrast, strongly enhanced photochemistry was seen by adding $LHCII_H$ to a mixture of QDs and RC_H (**Fig. 4.9A,** brown). As evidenced by the enhanced bleaching seen in a 5:1 mix of RC_H and QDs (**Fig. 4.9B**), some transfer of energy harvested by the QDs at 650 nm to RC_H will have occurred, but the major effect seen in the $LHCII_H$/RC_H/QD mixture is attributed to additional energy being passed from $LHCII_H$ to the RC_H, clearly showing that significant energy transfer could be switched on only when both proteins were modified with a His_{10}-tag and QDs were presented to act as an assembly hub (**Fig. 4.9C**).

The amount of P870-bleaching could be modulated by varying the $LHCII_H$:RC_H ratio in the mix used to form the conjugate, bleaching being strongest with two $LHCII_H$ for every RC_H and weakest with 0.5 $LHCII_H$ for every RC_H (**Fig. 4.9B**). This suggested that the overall energy transfer efficiency of the system could be adjusted through the size of the $LHCII_H$/QD antenna servicing each RC_H, which is resonant of natural photosystems where the number of light harvesting complexes per RC can be adjusted depending on factors such as incident light intensity[229].

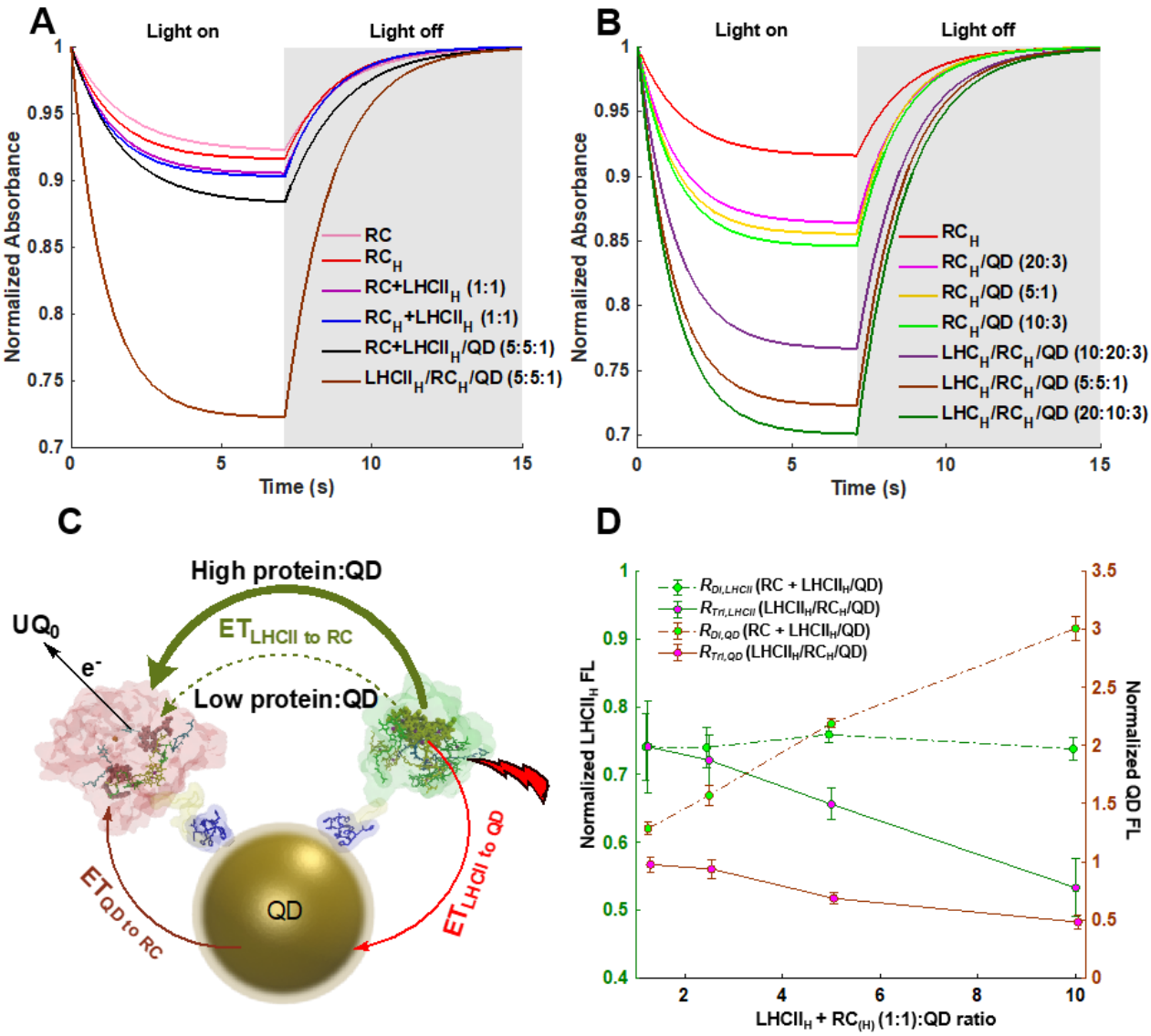


Figure 4.9. Energy transfer and charge separation in $LHCII_H/RC_H/QD$ conjugates. (A) Normalised RC photobleaching at 870 nm in mixes with native (RC) or His_{10}-tag modified (RC_H) complexes. Modest enhancements seen in samples other than $LHCII_H/RC_H/QD$ are likely to be due to some RC absorption of $LHCII_H$ or QD emission. (B) Normalised RC photobleaching at 870 nm in RC_H/QD or $LHCII_H/RC_H/QD$ conjugates with three different $LHCII_H$:RC_H ratios. For (A) and (B) excitation was at 651 nm for 7 seconds and data shown are fits to averaged kinetic traces (see **Fig. S4.4**). (C) Schematic of a $LHCII_H/RC_H/QD$ conjugate with routes of direct and indirect FRET. (D) Normalised $LHCII_H$ and QD emission as a function of the ratio of total-protein:QD (with $LHCII_H$:$RC_{(H)}$ = 1:1) for $LHCII_H/RC_H/QD$ conjugates or a composition-matched mixture of RC and $LHCII_H/QD$ conjugates. After deconvolution, $LHCII_H$ or QD emission was normalised relative to the emission from an equivalent concentration of $LHCII_H$ or QD in single component samples.

Intra-conjugate energy transfer was also apparent from a quenching of $LHCII_H$ emission that was seen when $LHCII_H$ and RC_H were mixed in the presence of QDs compared to the emission obtained when they were mixed, at the same concentrations, in the absence of QDs (**Fig. 4.10A**). Importantly, the extent of this quenching was the same irrespective of the order in which components were mixed, using three different addition sequences (**Fig. 4.10A**). This indicated that the state being measured was at equilibrium.

As depicted in **Fig. 4.9C**, Chl to BChl energy transfer in such a tri-component system might be expected to occur either by an indirect route when the QD hs appropriate optical properties (**Fig. 4.2A**), or additionally by a direct route when a high density of mixed $LHCII_H$ and RC_H are attached to a QD. To investigate this, QDs were titrated with increasing concentrations of an equimolar mixture of $LHCII_H$ and either RC or RC_H. Normalized $LHCII_H$ emission in a RC + $LHCII_H$/QD system ($R_{Di,LHCII}$) was fairly constant as the total-protein:QD ratio increased showing there was no reabsorption of $LHCII_H$ emission by RC proteins in solution, even at the highest concentration used (**Fig. 4.9D, green dash**). This was also the case when the $LHCII_H$:RC ratio was changed from 1:1 to 2:1 or 1:2 (**Fig. 4.10B**, green bars).

These data on a lack of variation in the level of $LHCII_H$ emission obtained when the composition of the RC + LHCIIH/QD system was changed were almost identical to a dataset obtained with a $LHCII_H$/$XylE_H$/QD control system, where $XylE_H$ can bind to the QD but is unable to receive energy from $LHCII_H$ (**Fig. 4.10B**, yellow bars). This demonstrated that the amount of $LHCII_H$ emission quenching was not affected by attachment of another optically-inactive protein to the QDs. In contrast to the invariance seen in the $LHCII_H$/QD + RC system (Fig. 4.9D, green dash), $LHCII_H$ emission in compositionally-matched $LHCII_H$/RC_H/QD samples ($R_{Tri,LHCII}$) declined as the total-protein:QD ratio increased (**Fig. 4.9D,** green solid). The difference between these trends, $LHCII_H$ quenching seen only when RCs were bound to QDs, provided a means to measure direct $LHCII_H \rightarrow RC_H$ energy transfer, the contribution of which became larger as the total-protein:QD ratio, and hence protein crowding, increased (see schematic in **Fig. 4.9C**).

Fig. 4.9D also shows normalised QD emission in the $LHCII_H$/QD + RC and $LHCII_H$/RC_H/QD systems ($R_{Di,QD}$ and $R_{Tri,QD}$ respectively). In the dataset from the system where the RC was not bound to the QD the amount of QD emission increased strongly as the total-protein:QD ratio increased (**Fig. 4.9D,** brown dash), consistent with increasing amounts of $LHCII_H \rightarrow$QD energy transfer. In contrast, when the RC_H was bound in a $LHCII_H$/RC_H/QD conjugate the level of QD

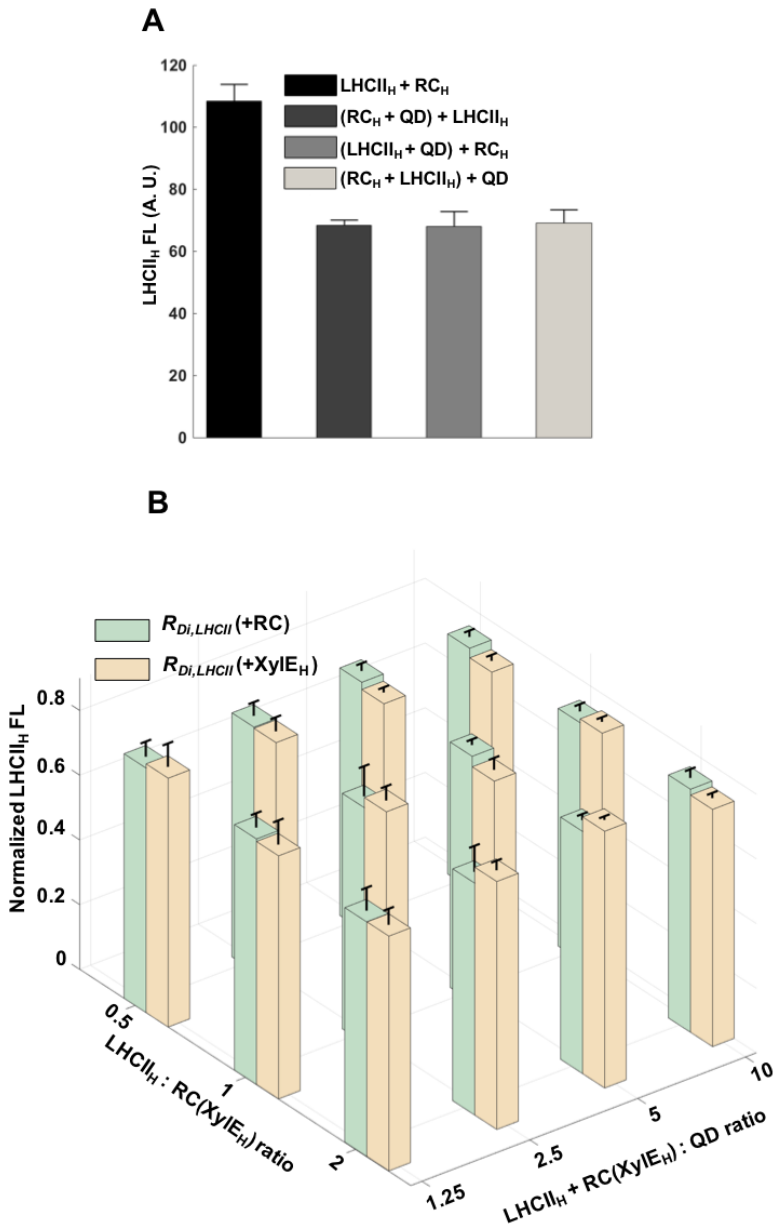


Figure 4.10. LHCII$_H$/RC$_H$/QD conjugate controls. (A) Quenching of LHCII$_H$ emission on its incorporation into a tri-component conjugate, relative to that seen on mixing with RC$_H$ in the absence of QDs. The initial two-component mix shown in brackets was supplemented with a third component. The final emission level for the three-component samples did not depend the addition order, indicating the system was at equilibrium. (B) LHCII$_H$ emission (*R*) normalized to a concentration-matched LHCII$_H$-only control as a function of LHCII$_H$:RC(XylE$_H$) ratio and total-protein:QD ratio. The level of LHCII emission was invariant, indicating no shading effect from RCs lacking a His-tag free in solution and no influence on LHCII$_H$ or QD quantum yield from attachment of a FRET-incapable protein XylE$_H$ to LHCII$_H$/QD conjugates.

emission declined gradually (**Fig. 4.9D,** brown solid), indicating that the RC_H was acting as an increasingly effective sink for energy being passed from $LHCII_H$. Similar trends in $LHCII_H$ and QD emission were seen in experiments with 2:1 and 1:2 mixes of $LHCII_H$ and either RC_H or native RCs (**Fig. S4.6A,B**).

4.2.7. Thermodynamic analysis of the composition of tri-component conjugates

The relative amounts of $LHCII_H$ and RC_H binding to a QD at different $LHCII_H$:RC_H and total-protein:QD ratios were measured by sucrose gradient ultracentrifugation, absorbance spectroscopy of gradient fractions and spectral deconvolution (**Fig. 4.11**). This showed that the relative amounts of bound $LHCII_H$ and RC_H was determined by the supplied concentrations up to total-protein:QD ratio of 10, but then biased toward RC_H at higher ratios. This bias can likely be attributed to subtle differences between the two proteins, particularly with regard to their size and the longer linker connecting the His_{10}-tag to the RC. The longer linker could mean that the RC_H has more freedom to move than $LHCII_H$ when tethered to the QD surface and is therefore better able to tolerate, or adjust to, competition from $LHCII_H$.

For a mixture of RC_H and $LHCII_H$ the kinetics of conjugate assembly can be viewed as a competitive process in which each protein binds to free "sites" on a QD that are not occluded by already-bound protein. For such a tri-component system there are four coordinates along which a conjugate associated with multiple copies of RC_H and $LHCII_H$ could migrate from defined state (**Fig. 4.12A**). Conjugates formed in protein mixtures with different total-protein:QD and $LHCII_H$:RC_H ratios were modelled using the reaction schemes:

$$LHCII_{His} + RC_{i-1}LHCII_{j-1}QD \underset{k_{r,LHCII}(i,j)}{\overset{k_{f,LHCII}(i,j)}{\rightleftharpoons}} RC_{i-1}LHCII_jQD \qquad (\mathbf{4.1})$$

$$RC_{His} + RC_{i-1}LHCII_{j-1}QD \underset{k_{r,RC}(i,j)}{\overset{k_{f,RC}(i,j)}{\rightleftharpoons}} RC_iLHCII_{j-1}QD \qquad (\mathbf{4.2})$$

where k_f and k_r are macroscopic association and dissociation rate constants, respectively, for each protein at given valencies (*i,j*) of the two proteins, and $LHCII_{His}$ and RC_{His} are the total protein population adjusted for any fraction unable to bind ($LHCII_{damage}$ = 5.7% (see Section 4.6.1); RC_{damage} = 4.1% (see Section 3.2.7).

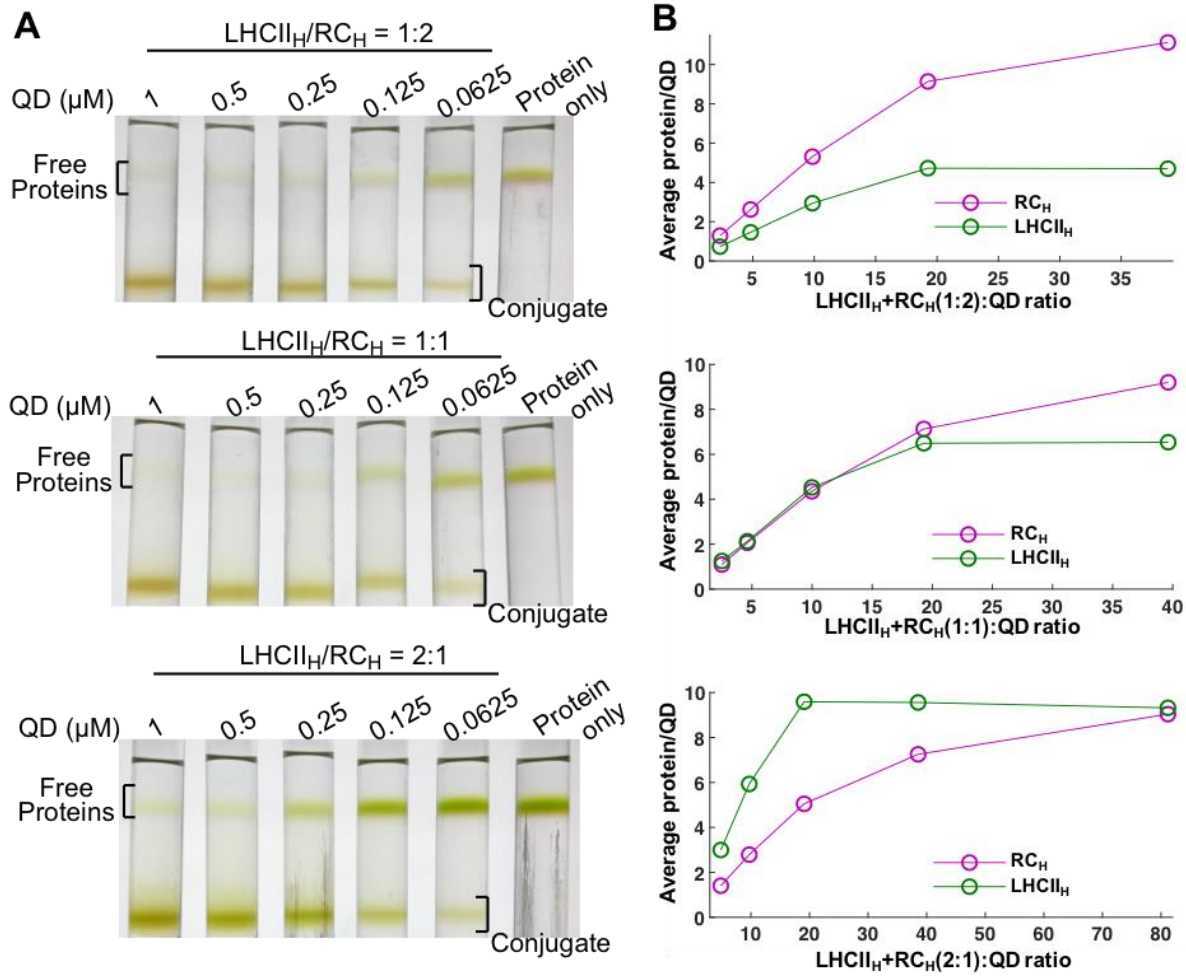


Figure 4.11. Variance of conjugate composition with initial protein mixture. (A) Separation of tri-component conjugates and unbound protein by ultracentrifugation on 0%/25%/60% sucrose density gradients. The total-protein concentration was fixed and the concentration of QDs varied to produce five total-protein:QD ratios. Three experiments were carried out in which the relative amounts of LHCII$_H$ and RC$_H$ in the protein mix were varied in (top to bottom) 1:2, 1:1 and 2:1 ratios. (B) Plot of average number of LHCII$_H$ (green) or RC$_H$ (purple) bound per QD against total-protein:QD ratio for the three LHCII$_H$:RC$_H$ ratios. Values were obtained by recording absorbance spectra from deconstructed gradients and deconvoluting the LHCII$_H$ and RC$_H$ contributions.

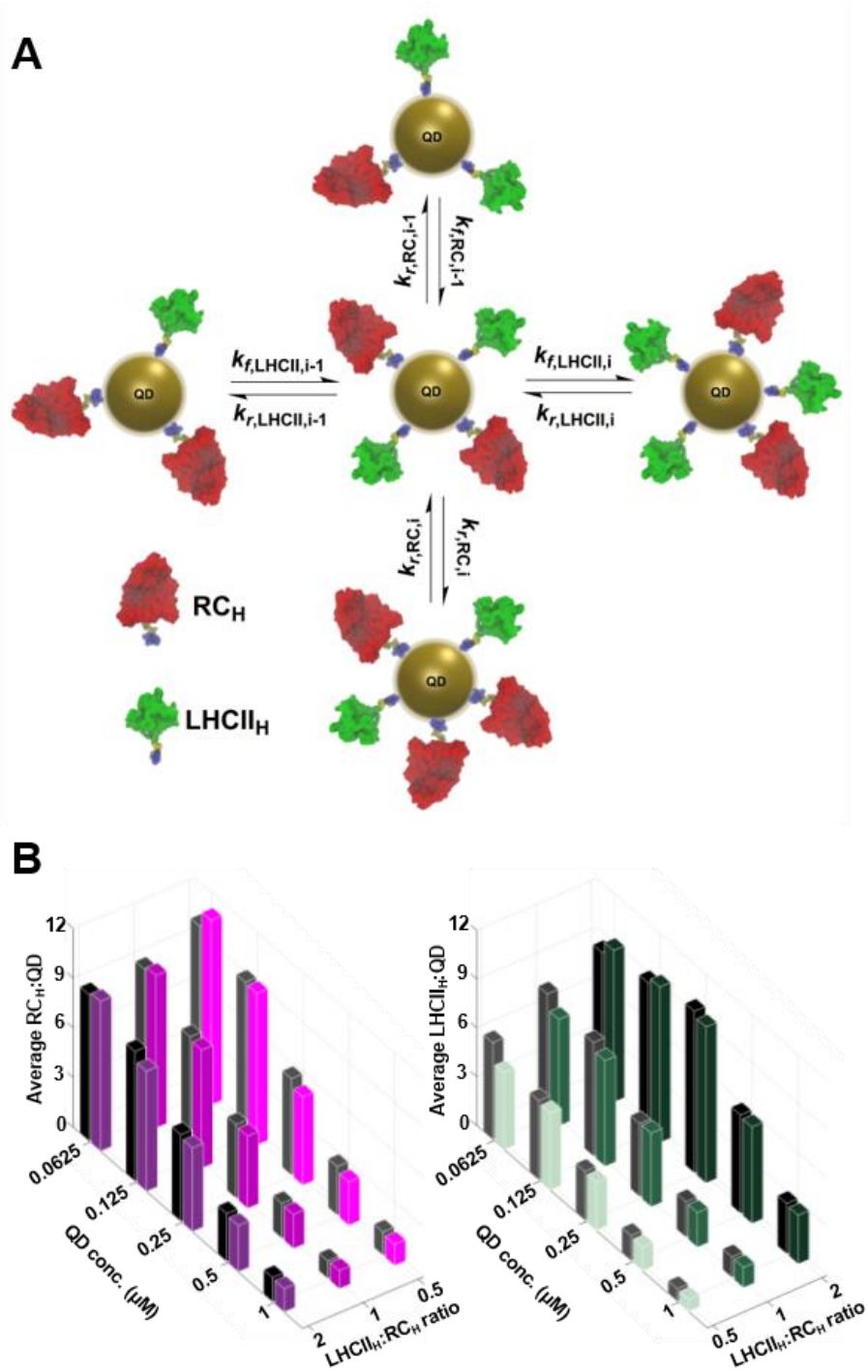


Figure 4.12. Modelling of $LHCII_H$/RC_H/QD conjugate assembly. (A) Scheme for association/dissociation of RC_H (red) and $LHCII_H$ (green) with the surface of a QD as described by **Eqs. 4.1** and **4.2.** (B) Comparison of data for (colour), and best-fit simulation of (greyscale), the average RC_H:QD (left) and $LHCII_H$:QD (right) composition of conjugates as a function of total-protein:QD ratio (achieved by varying [QD]) for three different $LHCII_H$:RC_H ratios.

The macroscopic association rate constants ($k_{f,RC}$ *and* $k_{f,LHCII}$) in **Eqs. 4.1** and **4.2** depend on the availability of free QD surface at which to bind. This availability was parameterised by score matrices of fractional occupation of a QD surface by either RC_H (S_{RC}) or $LHCII_H$ (S_{LHCII}) in tri-component conjugates, that were constructed using:

$$S_{RC}(i,j) = \frac{i}{\bar{n}_{RC}} + \alpha\frac{j}{\bar{n}_{LHCII}} \quad \textbf{(4.3)}$$

$$S_{LHCII}(i,j) = \beta\frac{i}{\bar{n}_{RC}} + \frac{j}{\bar{n}_{LHCII}} \quad \textbf{(4.4)}$$

where $\bar{n}_{RC}$ and $\bar{n}_{LHCII}$ were maximum permitted binding valencies determined from studies of di-component conjugates (15 and 17, respectively – see Sections 3.2.7 and 4.2.3). Parameter *i* denoted any possible RC_H:QD valency between 0 and 15 and *j* denoted any possible $LHCII_H$:QD valency between 0 and 17.

Parameter α was included in **Eq. 4.3** to account for the influence of already bound $LHCII_H$ on binding of an additional RC_H and parameter β was included in **Eq. 4.4** to account for the reverse influence; these were required to account for the experimental observation that binding biased toward RC_H at total-protein:QD ratios >10.

The macroscopic association (k_f) and dissociation (k_r) rate constants for RC_H and $LHCII_H$ shown in **Eqs. 4.1** and **4.2** were deduced from pairs of imaginary microscopic kinetic constants (k_{on}, k_{off}) that were determined from the thermodynamic constants $k_{micro,RC}$ = 8.2 μM^{-1} (from Chapter 3) and $k_{micro,LHCII}$ = 7.4 μM^{-1} (this chapter), following the relationship $k_{micro} = k_{on}/k_{off}$. Matrices of k_f *and* k_r for each protein were determined from:

$$k_{f,RC}(i,j) = \bar{n}_{RC}(1 - S_{RC}(i,j))k_{on,RC} \quad \textbf{(4.5)}$$

$$k_{f,LHCII}(i,j) = \bar{n}_{LHCII}(1 - S_{LHCII}(i,j))k_{on,LHCII} \quad \textbf{(4.6)}$$

$$k_{r,RC}(i,j) = ik_{off,RC} \quad \textbf{(4.7)}$$

$$k_{r,LHCII}(i,j) = jk_{off,LHCII} \quad \textbf{(4.8)}$$

where *i* = 1, 2,…, $\bar{n}_{RC}$ and *j* = 1, 2,…, $\bar{n}_{LHCII}$.

The model summarised in **Eqs. 4.1-4.8** was used to simulate the experimental data (**Fig. 4.11B**) on average RC_H:QD and $LHCII_H$:QD stoichiometries for $LHCII_H/RC_H$/QD conjugates formed in mixtures with three $LHCII_H$:RC_H ratios and (for each) five total-protein:QD ratios. To be relevant to the binding experiments this was carried out by fixing [$LHCII_H$+RC_H] at 2.5 μM while varying [QD]. **Fig. 4.12B** shows the best fit of the simulation (black/grey bars) to the data for RC_H (purple/pink bars) and $LHCII_H$ (green/cyan bars), which was achieved with α = 0.63 and β = 0.78. That both these parameters had values less than one suggested that both $LHCII_H$ and RC_H had a smaller blocking effect on binding of the partner protein than on themselves, and the smaller value for β was consistent with RC_H being more competitive for binding than $LHCII_H$ despite their very similar micro kinetic constants. The S_{RC} and S_{LHCII} score matrices for fractional occupancy from this best fit are shown in **Fig. 4.13** with all values ≤ 1 highlighted in colour. Only these permitted fractional occupancies were considered as a value larger than one corresponded to an overloaded $LHCII_H/RC_H$/QD configuration. The associated k_f *and* k_r matrices for each protein are shown in **Figs. 4.13B**,**C**.

Population distributions around the average compositions of the tri-component conjugates were then modelled. An example result for conjugates formed from an initial 5:5:1 $LHCII_H$:RC_H:QD mix is shown in **Fig. 4.14A**. The most favoured compositions were 4 $LHCII_H$+5 RC_H, 5 $LHCII_H$+4 RC_H and 4 $LHCII_H$+4 RC_H close to the supplied composition, with a broad distribution around these values. Population distributions for six other total-protein:QD ratios with equal molar $LHCII_H$ and RC_H are shown in **Fig. S4.7**.

Modelling of the average RC_H:QD or $LHCII_H$:QD ratio for a broader range of total-protein:QD ratios (1.25 to 128) and solution $LHCII_H$:RC_H ratios (0.3 to 2.2) showed that binding of RC_H was relatively insensitive to the presence of $LHCII_H$ at high total-protein:QD ratios **(Fig. 4.14B)**. However, when this ratio was below approximately 13, the bound fractions were directly determined by the ratio of the two proteins in solution (**Fig. 4.14B**), in agreement with experimental observations (**Fig. 4.11B**).

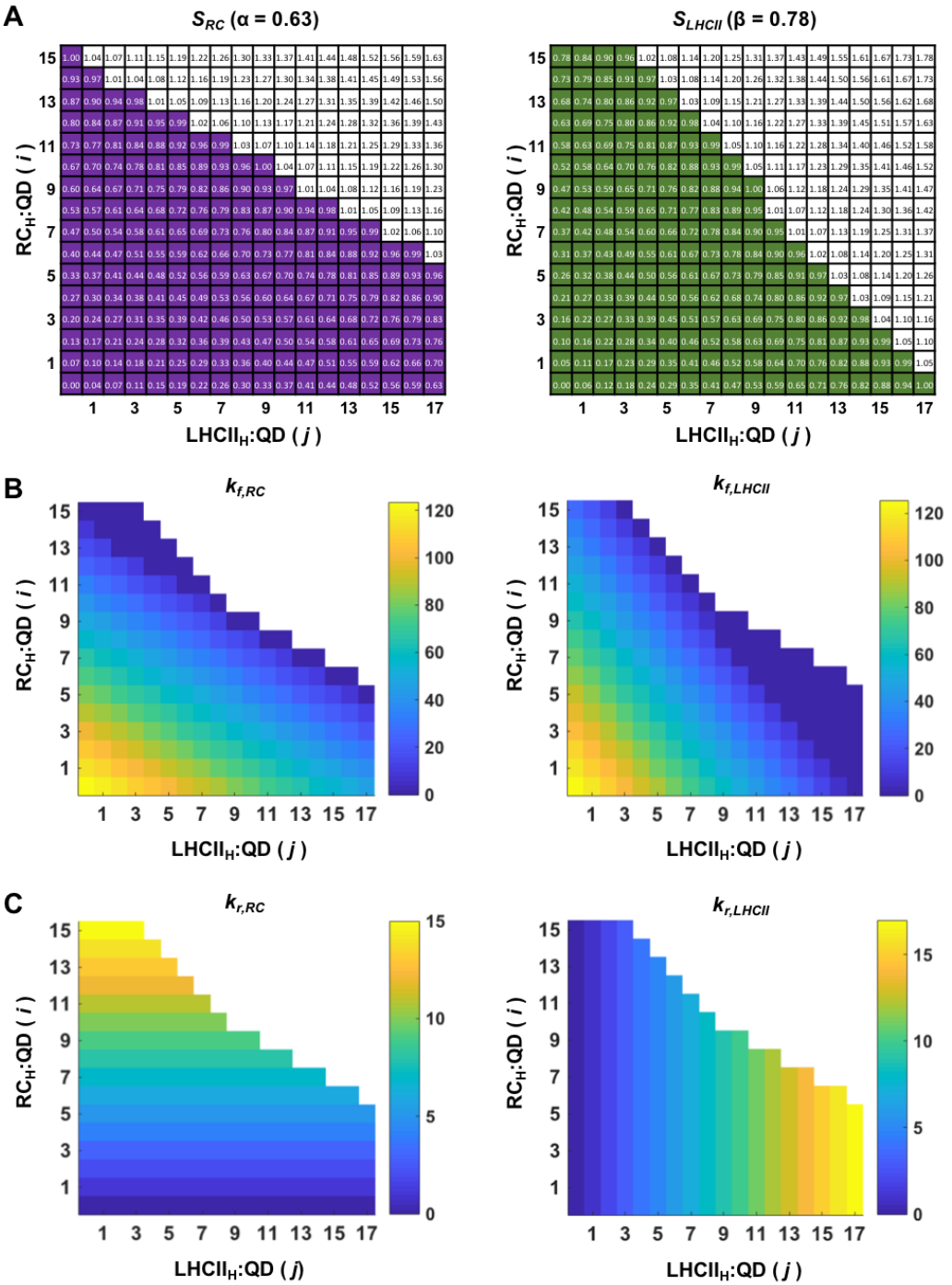


Figure 4.13. Parameterization of LHCII$_H$/RC$_H$/QD conjugation modelling. (A) Fractional occupancy of a QD surface by protein (*S*) at all possible RC$_H$:QD and LHCII$_H$:QD ratios calculated using **Eq. 4.3** for S_{RC} (left) and **Eq. 4.4** for S_{LHCII} (right). (B) Assembly rate constants (k_f) calculated using **Eq. 4.5** for RC (left) and **Eq. 4.6** for LHCII (right) at all possible RC$_H$:QD and LHCII$_H$:QD ratios. (C) Disassembly rate constants (k_r) calculated using **Eq. 4.7** for RC (left) and **Eq. 4.8** for LHCII (right) LHCII at all possible RC$_H$:QD and LHCII$_H$:QD ratios.

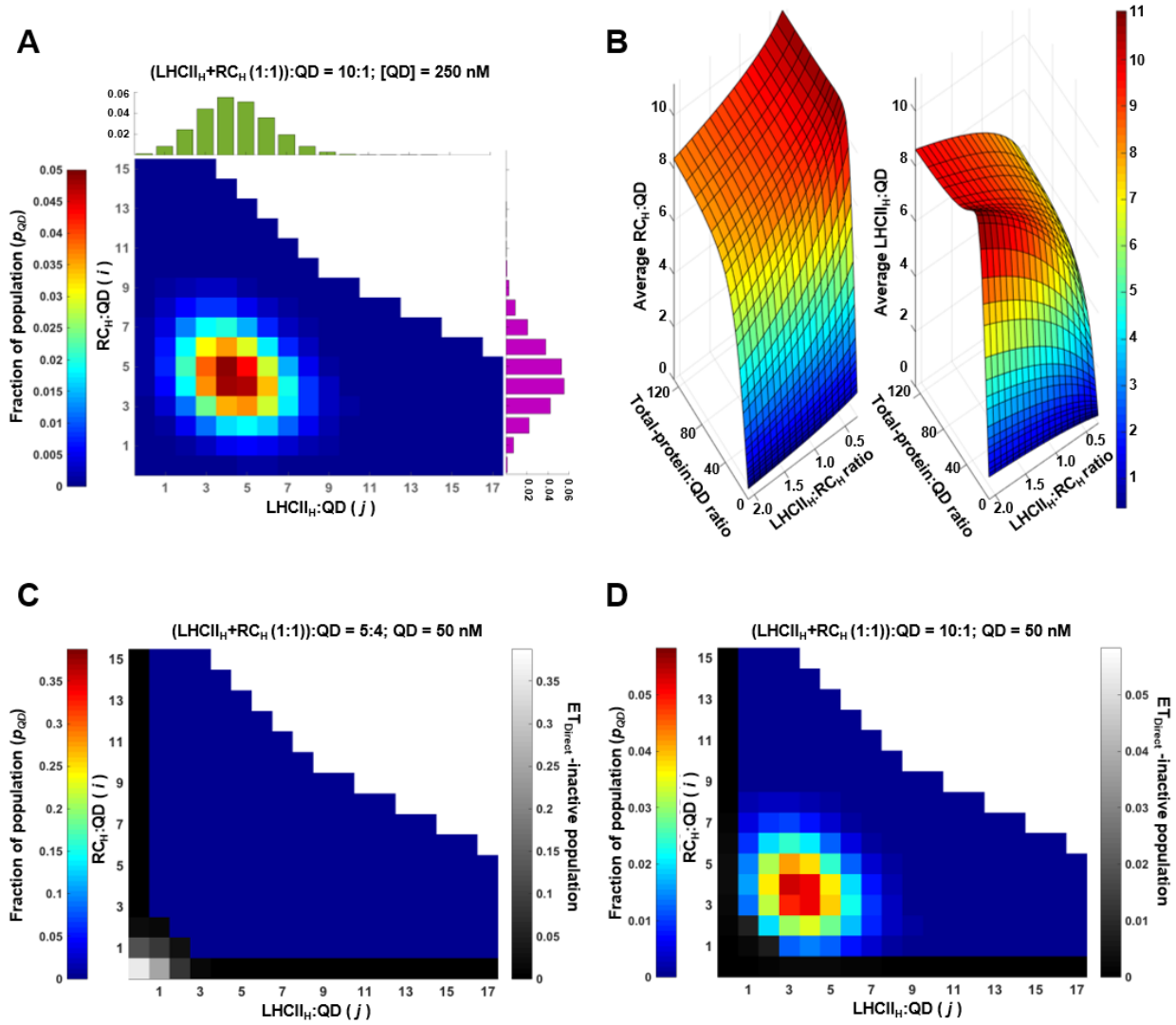


Figure 4.14. Extents of LHCII$_H$/RC$_H$/QD heterogeneity and FRET-inactive sub-populations. (A) Computed population distribution for conjugates formed from a 5:5:1 mix of RC$_H$, LHCII$_H$ and QDs. Individual distributions of LHCII$_H$:QD and RC$_H$:QD are shown with green and purple bars, respectively. (B) Continuous plots of modelled average RC$_H$:QD (left) and LHCII$_H$:QD (right) in conjugates as a function of total-protein:QD ratio between 0.625 to 128, and for LHCII$_H$:RC$_H$ ratios between 0.3 and 2.2. Each data point was the average RC$_H$:QD or LHCII$_H$:QD ratios by summing all bound RC$_H$ or LHCII$_H$ to QD obtained from simulation equilibria as shown in (A). (C,D) Conjugate compositions in mixes with total-protein:QD ratios of 1.25:1 (C) and 10:1 (D) with LHCII$_H$:RC$_H$ = 1:1. Conjugate compositions that are inactive for FRET are shown in greyscale, and include all conjugates where the number of RC$_H$ or LHCII$_H$ bound to a QD is zero (left column and bottom row within each plot).

4.2.8. Population distribution and energy transfer

An insight from this distribution modelling was that, in a mixture of composition $LHCII_H$:RC_H:QD = 5:5:8, nearly 84 % of QDs either bound no protein or bound only RC_H or only $LHCII_H$, and thus were predicted not to be able to carry out $LHCII_H \rightarrow RC_H$ ET. This is illustrated in the distribution map in **Fig. 4.14C**, where the compositions incapable of producing detectable $LHCII_H \rightarrow RC_H$ ET are shown in greyscale. This agreed with fluorescence titration experiments (**Fig. 4.9D**) where, at the equivalent total-protein:QD ratio of 1.25, equal amounts of $LHCII_H$ emission were seen in RC+$LHCII_H$/QD and $LHCII_H$/RC_H/QD conjugates (coincidence of green-dash and green-solid data on left of Fig. 4.9D).

As the $LHCII_H$:RC_H:QD composition rose to 5:5:1 (**Fig. 4.14D** - and see **Fig. S4.8D-G** for two intermediate cases) the fraction of the population incapable of $LHCII_H \rightarrow RC_H$ ET shrank as the numbers of QDs with multiple bound RC_H and $LHCII_H$ increased. Again, this was reflected in the data from fluorescence experiments (**Fig. 4.9D**) where, at the equivalent total-protein:QD ratio of 10, substantially more quenching of $LHCII_H$ emission were seen in $LHCII_H$/RC_H/QD conjugates compared to RC+$LHCII_H$/QD conjugates (divergence of green-dash and green-solid data on left of Fig. 4.9D). When the ratio of $LHCII_H$:RC_H moved away from unity this introduced bias in the population distribution that tended to slightly increase the FRET-inactive sub-population especially at low total-protein:QD ratios (depicted in **Fig. S4.8A-D** and **H-K**), leading to a reduction of energy flow.

4.2.9. Estimation of the efficiencies of direct and indirect energy transfer.

The efficiencies of direct (E_{Direct}) and indirect ($E_{Indirect}$) FRET between $LHCII_H$ and RC_H in the tri-component conjugates were estimated using the data on $LHCII_H$ and QD emission in $LHCII_H$/QD conjugates and $LHCII_H$/RC_H/QD conjugates with three different $LHCII_H$:RC_H ratios (i.e. the datasets shown in **Figs. 4.9D** and **S4.6**).

4.2.9.1. Estimation of E_{Direct}.

E_{Direct} was estimated from the relative levels of $LHCII_H$ emission in the tri-component and di-component conjugates. To do this it was necessary to determine parameters $F_{Tri,LHCII}$ and $F_{Di,LHCII}$ which were the fraction (F) of supplied $LHCII_H$ bound to the tri- and di-component conjugates. These were determined from **Eq. S4.3**:

$$F_{Tri(Di),LHCII} = \frac{[QD]_T}{[LHCII]_T} \sum_{i=0}^{\bar{n}_{RC}} \sum_{j=0}^{\bar{n}_{LHCII}} jp_{QD}(i,j) \quad \textbf{(4.9)}$$

where *i* and *j* were the possible stoichiometries of RC_H and $LHCII_H$ to QD in a given conjugate and $p_{QD}(i,j)$ was the population distribution of that conjugate (depicted in **Fig. S4.7**). $[LHCII]_T$ and $[QD]_T$ were the total supplied concentration of $LHCII_H$ and QD. The deduced bound fraction from **Eq. 4.9** was exactly the same as the difference between total $LHCII_H$ or QD and the unbound fraction, indicating that the setup was correct and that the simulation was stable with the chosen ordinary differential equation solution.

To obtain E_{Direct} it was also necessary to quantify the relative levels of $LHCII_H$ emission in the tri-component and di-component conjugates. Parameter $R_{Tri,LHCII}$ represented the $LHCII_H$ emission in tri-component $LHCII_H/RC_H$/QD conjugates normalized to concentration-matched $LHCII_H$-only controls, whilst $R_{Di,LHCII}$ was the corresponding parameter for a compositionally-equivalent mixture of $LHCII_H$/QD di-component conjugates in the presence of non-His-tagged RCs. Corrections for an unbound protein fraction were implemented as the calculation of efficiency is to total QD. On the premise that the observed energy transfer is the result of a competition between different kinetics[197], $R_{Tri,LHCII}$ and $R_{Di,LHCII}$ therefore represented the relationships of these exciton relaxation pathways according to.

$$\frac{1 - R_{Di,LHCII}}{F_{Di,LHCII}} = \frac{k_{LHCIItoQD}}{k_{others} + k_{FL} + k_{LHCIItoQD}} \quad \textbf{(4.10)}$$

$$\frac{1 - R_{Tri,LHCII}}{F_{Tri,LHCII}} = \frac{k_{LHCIItoQD} + k_{Direct}}{k_{others} + k_{FL} + k_{LHCIItoQD} + k_{Direct}} \quad \textbf{(4.11)}$$

where k_{others} is the rate of all non-fluorescent relaxation pathways, k_{FL} is the fluorescent pathway, k_{Direct} is direct energy transfer and $k_{LHCIItoQD}$ is energy transfer from $LHCII_H$ to QD as $k_{LHCIItoQD}$.

E_{Direct} was estimated using $R_{Di,LHCII}$, $R_{Tri,LHCII}$, $F_{Di,LHCII}$ and $F_{Tri,LHCII}$ as shown in **Eq. 4.12**.

$$E_{Direct} = \frac{k_{Direct}}{k_{others} + k_{FL} + k_{LHCIItoQD} + k_{Direct}} = 1 - \frac{1 - \frac{1-R_{Tri,LHCII}}{F_{Tri,LHCII}}}{1 - \frac{1-R_{Di,LHCII}}{F_{Di,LHCII}}} \quad \textbf{(4.12)}$$

Rearrangement of **Eq. 4.12** produced **Eq. 4.13**, which was used to determine E_{Direct}.

$$E_{Direct} = \frac{F_{Di,LHCII}(1 - R_{Tri,LHCII}) - F_{Tri,LHCII}(1 - R_{Di,LHCII})}{F_{Tri,LHCII}(F_{Di,LHCII} - 1 + R_{Di,LHCII})} \quad \textbf{(4.13)}$$

Having determined E_{Direct}, the efficiency of energy transfer from $LHCII_H$ to the QD could then be determined by:

$$E_{LHCIItoQD} = \frac{k_{LHCIItoQD}}{k_{others} + k_{FL} + k_{LHCIItoQD} + k_{Direct}} = \frac{1 - R_{Tri,LHCII}}{F_{Tri,LHCII}} - E_{Direct} \quad \textbf{(4.14)}$$

It is obvious that there would be expected to be competition between E_{Direct} and $E_{LHCIItoQD}$. Generally speaking, $E_{LHCIItoQD}$ dominated the energy flow in the low total-protein:QD regime. while the E_{Direct} became increasingly competitive at high total-protein:QD ratios, causing $E_{LHCIItoQD}$ to decline.

4.2.9.2. Estimation of $E_{Indirect}$ – effect of fraction of FRET-inactive conjugates.

For the indirect route, $E_{Indirect}$ was estimated using **Eq. 4.15** from the effective efficiencies of $LHCII_H$→QD energy transfer ($E'_{LHCII\ to\ QD}$) and QD→RC_H energy transfer ($E'_{QD\ to\ RC}$), considering only those conjugates within any given population that had both $LHCII_H$ and RC_H associated.

$$E_{Indirect} = E'_{LHCIItoQD} E'_{QDtoRC} \quad \textbf{(4.15)}$$

$E'_{LHCIItoQD}$ and E'_{QDtoRC} were adjusted effective energy transfer efficiencies based on measured values of $E_{LHCIItoQD}$ and E_{QDtoRC}. The reason for using effective efficiencies was that indirect energy transfer required both $LHCII_H$ and RC_H to be bound to a QD, and the proportion of the conjugate population where this was not the case grew as the total-protein:QD ratio became lower. This is illustrated in **Fig. 4.15**, where the populations lacking either one $LHCII_H$ or one RC_H are highlighted for two types of tri-component conjugate. This fraction is small when the total-protein:QD ratio is high (**Fig. 4.15A**) but becomes large when the total-protein:QD ratio is small (**Fig. 4.15B**).

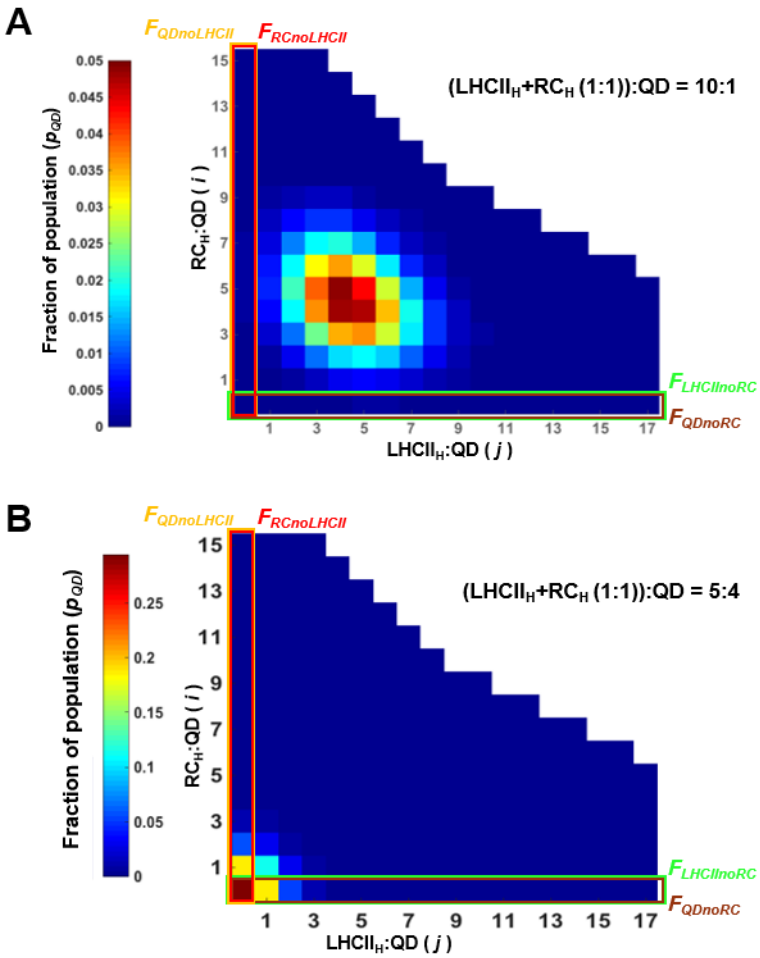


Figure 4.15. Highlight of fraction of conjugates incapable of indirect energy transfer. The light and dark brown boxes denote the conjugates with only RC_H or $LHCII_H$. (A) Simulation end-state of protein:QD = 10:1 and [QD] = 100nM. (B) Simulation end-state of protein:QD = 5:4 and [QD] = 100nM. To note, the fraction values of the same set of conjugates were slightly different for bound RC_H (red) and $LHCII_H$ (green), as was seen from **Eqs. 4.16** and **4.17**

The fraction (*F*) of $LHCII_H$ or QDs with no associated RC_H could be estimated from:

$$F_{LHCIInoRC} = \frac{[QD]_T}{[LHCII]_T F_{Tri,LHCII}} \sum_{j=0}^{\bar{n}_{LHCII}} j p_{QD}(0,j) \quad \textbf{(4.16)}$$

$$F_{QDnoRC} = \sum_{j=0}^{\bar{n}_{LHCII}} p_{QD}(0,j) \quad \textbf{(4.17)}$$

in which population subfractions without RC_H ($p_{QD}(0,j)$) were summed to calculate the "no RC" population ($F_{LHCIInoRC}$ and F_{QDnoRC} to bound [LHCII] and total [QD]).

4.2.9.3. Estimation of $E_{Indirect}$ – determination of $E'_{LHCIItoQD}$.

Since energy transfer between a $LHCII_H$ and a QD could be explained by a multiple donor – single acceptor FRET scheme, the amount of energy transfer was directly related to the donor concentration (see **Section 4.6.1**). Therefore, the effective $E_{LHCIItoQD}$ (denoted $E'_{LHCIItoQD}$) could be deduced from:

$$E'_{LHCIItoQD} = E_{LHCIItoQD}(1 - F_{LHCIInoRC}) \quad \textbf{(4.18)}$$

Values of $E'_{LHCIItoQD}$ determined for four total-protein:QD ratios for each of three $LHCII_H$:RC_H ratios are summarized in **Fig. 4.16A** (red bars). These were used in **Eq. 4.15** to calculate $E_{Indirect}$.

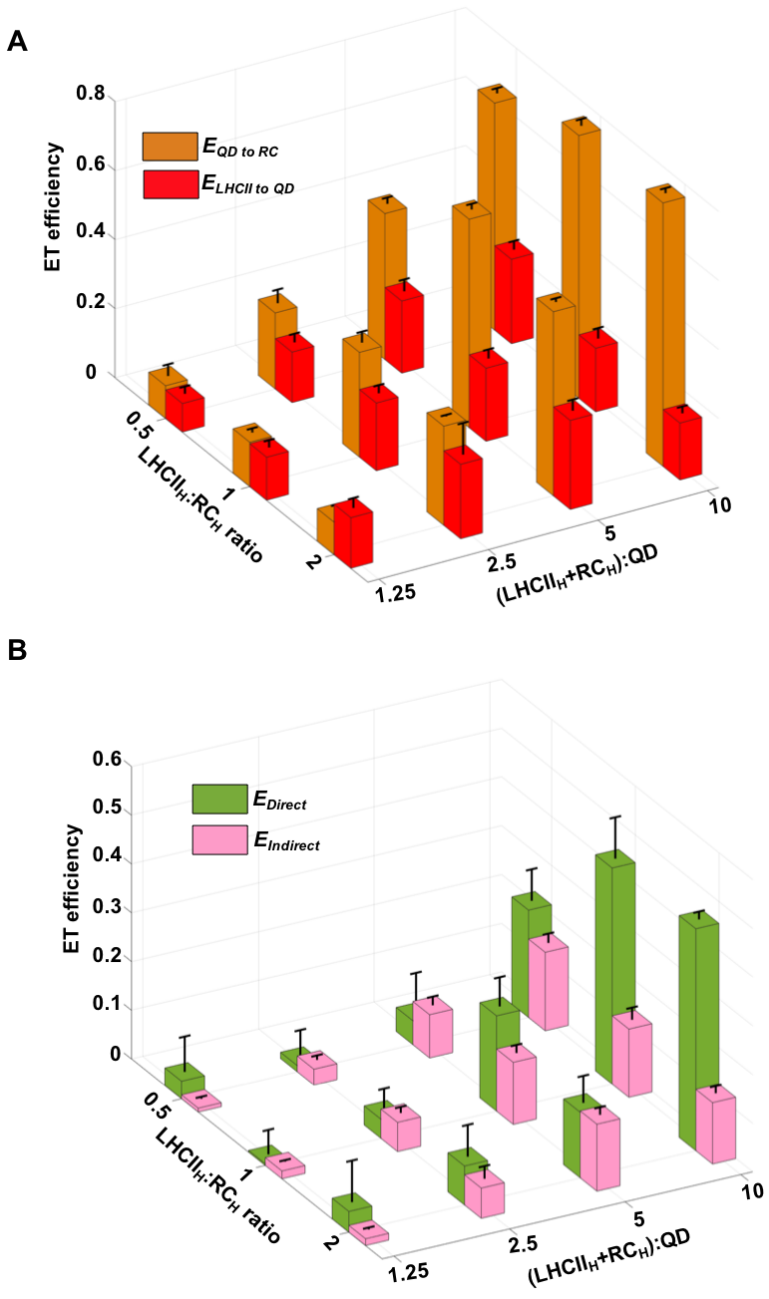


Figure 4.16. Energy transfer efficiencies at all tested $LHCII_H$:RC_H:QD ratios. (A) **E**fficiencies for $LHCII_H$ to QD and QD to RC_H energy transfer for all assessed compositions. The efficiency of E'_{QDtoRC} increased with more bound RC_H, while a less drastic increase in efficiency of $E'_{LHCIItoQD}$ with increasing total-protein:QD and $LHCII_H$:RC_H ratio due to the high FRET incapable fraction at low total-protein:QD ratios and increasing energy competition at high total-protein:QD. (B) Efficiencies of direct and indirect energy transfer for all assessed combinations, where $E_{Indirect}$ was the multiple product of $E'_{LHCIItoQD}$ and E'_{QDtoRC} shown in (A).

4.2.9.4. Estimation of $E_{Indirect}$ – determination of E'_{QDtoRC}.

The second piece of information required to determine $E_{Indirect}$ was the effective energy transfer efficiency from the QD to attached RC_H (E'_{QDtoRC}), which could be extracted from QD emission in the di-component mixture $LHCII_H$ + RC ($R_{Di,QD}$) and the corresponding tri-component mixture $LHCII_H$ + RC_H ($R_{Tri,QD}$), both normalized to the emission from concentration-matched QD-only samples. As described previously, there was a competition for energy transfer from $LHCII_H$ that affected the entry of energy to QDs with or without the presence of RC_H. Therefore, the parameter $R_{Di,QD}$ was adjusted for the difference in the amount of energy entry using:

$$R'_{Di,QD} = (R_{Di,QD} - 1)\frac{E_{LHCIItoQD} F_{Tri,LHCII}}{1 - R_{Di,LHCII}} + 1 \qquad (\mathbf{4.19})$$

Here $R'_{Di,QD}$ was $R_{Di,QD}$ corrected for less additional energy input from $LHCII_H$ as the result of competition from bound RC_H. In fact, a difference could only be seen at high total-protein:QD ratios (10:1) where E_{Direct} surpassed $E_{LHCIItoQD}$.

Having obtained $R'_{Di,QD}$ the overall energy transfer efficiency from the QDs to bound RC_H (E_{QDtoRC}) was calculated using:

$$E_{QDtoRC} = 1 - \frac{R_{Tri,QD}}{R'_{Di,QD}} \qquad (\mathbf{4.20})$$

As was discussed in Chapter 3 for RC_H/QD conjugates, the overall efficiency of energy transfer from a QD to bound RC_H could be explained by a multiple acceptor – single donor FRET scenario assuming a constant energy transfer efficiency ($E_{DA,RC}$) from the QD to every bound RC_H. Based on this, the effective energy transfer efficiency E_{QDtoRC} (denoted E'_{QDtoRC}) could be derived with information on the population distribution.

First, the relative emission of conjugates depending on occupation by $LHCII_H$ had to be estimated across all possible $LHCII_H$/RC_H/QD configurations to calculate the contribution of every configuration to the observed average apparent value. According to the measured energy transfer efficiency from LHCII to QD ($E_{LHCIItoQD}$) and the measured molar $LHCII_H$ and QD absorption ratio (Abs_{LHCII}/Abs_{QD}), the simulated QD normalized emission therefore can be obtained by:

$$R_{sim,QD} = \sum_{i=0}^{\bar{n}_{RC}} \sum_{j=0}^{\bar{n}_{LHCII}} p_{QD}(i,j)(1 + jE_{LHCIItoQD}\frac{Abs_{LHCII}}{Abs_{QD}}) \qquad (\mathbf{4.21})$$

where $R_{sim,QD}$ represented the normalized QD fluorescence from simulation, values of which were consistent with experimental observations ($R'_{Di,QD}$). Then, the relationship between $E_{DA,RC}$ and E_{QDtoRC} could be deduced by summing all contributions of bound RC_H to $LHCII_H$/QD conjugates at each corresponding stoichiometry using:

$$E_{QDtoRC} = \frac{1}{R_{sim,QD}} \sum_{i=0}^{\bar{n}_{RC}} \sum_{j=0}^{\bar{n}_{LHCII}} p_{QD}(i,j) \frac{iE_{DA,RC}}{(i-1)E_{DA,RC}+1} (1 + jE_{LHCIItoQD} \frac{Abs_{LHCII}}{Abs_{QD}}) \qquad \textbf{(4.22)}$$

Comparison of E_{QDtoRC} with $E_{DA,RC}$ revealed a non-linear relationship (**Fig. 4.17A-D**), the non-linearity increasing with more RC_H (acceptors) per QD (donor). The deduced $E_{DA,RC}$ had a mean of 0.52 ± 0.04 across all tested $LHCII_H$:RC_H:QD ratios, consistent with the value of 0.53 ± 0.02 in Chapter 3 for RC/QD conjugates with a RC_H:QD = 5:8.

Finally, the effective energy transfer efficiency from QDs to RC_H (E'_{QDtoRC}) was calculated from E_{QDtoRC} by removing the proportion of RC_H/QD without an $LHCII_H$ attached (j = 0):

$$E'_{QDtoRC} = E_{QDtoRC} - \frac{1}{R_{sim,QD}} \sum_{i=0}^{\bar{n}_{RC}} p_{QD}(i,0) \frac{iE_{DA,RC}}{(i-1)E_{DA,RC}+1} \qquad \textbf{(4.23)}$$

In fact, E'_{QDtoRC} and E_{QDtoRC} seemed to have a linear relationship as shown in **Fig. 4.17E-H** and the most loss of efficiency was seen at low total-protein:QD ratios because in this regime the predominant population of conjugates had only either RC_H or $LHCII_H$, which was also the reason for the low effective $LHCII_H$ to QD efficiency ($E'_{LHCIItoQD}$) (**Fig. 4.16A**).

Computed values for the effective QD to RC_H ET efficiency (E'_{QDtoRC}) are shown in **Fig. 4.16A.** These were used for the calculation of $E_{Indirect}$ by combining with $E'_{LHCIItoQD}$ using **Eq. 4.15**.

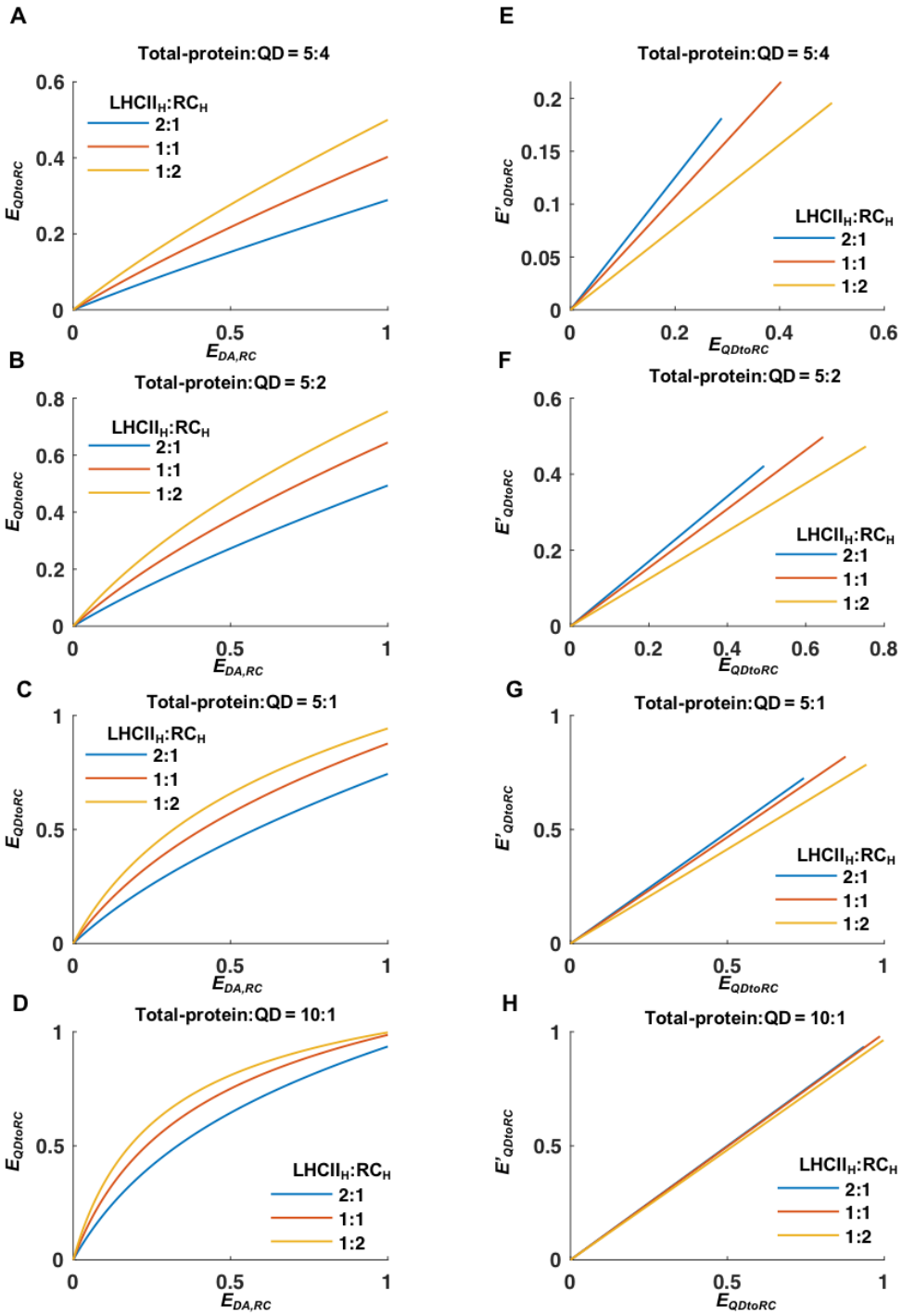


Figure 4.17. Correlations between E'_{QDtoRC} and $E_{DA,RC}$ and between E'_{QDtoRC} and E'_{QDtoRC}. (A-D) Nonlinear correlation between E'_{QDtoRC} and the single donor-acceptor $E_{DA,RC}$ for three LHCII$_H$:RC$_H$ ratios, calculated with **Eq. 4.22** for four total-protein:QD ratios. (E-H) Linear relationship between E'_{QDtoRC} and E_{QDtoRC} for three LHCII$_H$:RC$_H$ ratios, calculated with **Eq. 4.23** for four total-protein:QD ratios.

4.2.10. Relative contributions of direct and indirect $LHCII_H \rightarrow RC_H$ energy transfer in tri-component conjugates

Variations of E_{Direct} and $E_{Indirect}$ with total-protein:QD ratio at a $LHCII_H$:RC_H equal to one are compared in **Fig. 4.18A**. Both $E_{Indirect}$ and E_{Direct} were negligible at $LHCII_H$:QD or RC_H:QD ratios less than one but rose markedly up to ten total-proteins per QD. Indirect energy transfer switched on before the direct pathway since energy flow mediated by QDs did not require apposition of individual $LHCII_H$ and RC_H molecules for direct energy transfer, while E_{Direct} outstripped $E_{Indirect}$ at high total-protein:QD ratios because of shortening of average $LHCII_H$ – RC_H separation. Meanwhile, multiple $LHCII_H$ donor/multiple RC_H acceptor FRET interactions enhanced the apparent energy transfer efficiency. Similar trends were also obtained for $LHCII_H$:RC_H ratios other than one, as depicted in **Fig. 4.16B**.

For any given total-protein:QD ratio, the relative contributions of E_{Direct} and $E_{Indirect}$ to overall efficiency ($E_{Overall}$) were sensitive to the $LHCII_H$:RC_H ratio (**Fig. 4.18D**), with the overall energy transfer efficiency ($E_{Overall}$) seeming maximize at an $LHCII_H$:RC_H ratio equal to one (**Fig. 4.18D, maroon**). This can be envisaged as a situation where, on average, every bound $LHCII_H$ was adjacent to multiple RC_H and *vice versa*, producing an E_{Direct} that was about double that of $E_{Indirect}$. When the $LHCII_H$:RC_H ratio was increased to two (**Fig. 4.13D**, green), a decrease in $E_{Overall}$ was obtained due to a drop in E_{Direct}. This could be due to a sub-fraction of the major $LHCII_H$ population not being sufficiently close to a member of the minor RC_H population for effective direct FRET to take place, causing a waste of absorbed energy. However, for every RC_H there were more $LHCII_H$ within FRET distance and therefore enhanced P870 photobleaching was observed (**Fig. 4.9B**, dark green). In contrast, when RC_H was supplied in a 2:1 ratio over $LHCII_H$ the ET efficiencies did not deviate substantially from those seen with a 1:1 mix (**Fig. 4.18D**, purple). In this case, on average, each $LHCII_H$ will be adjacent to multiple RC_H, enabling efficient ET. However, this scenario also has the possibility of a fraction of RC_H that are not adjacent to an $LHCII_H$ and so do not receive energy efficiently, accounting for the observed lower level of P870 photobleaching (**Fig. 4.9B**, dark purple).

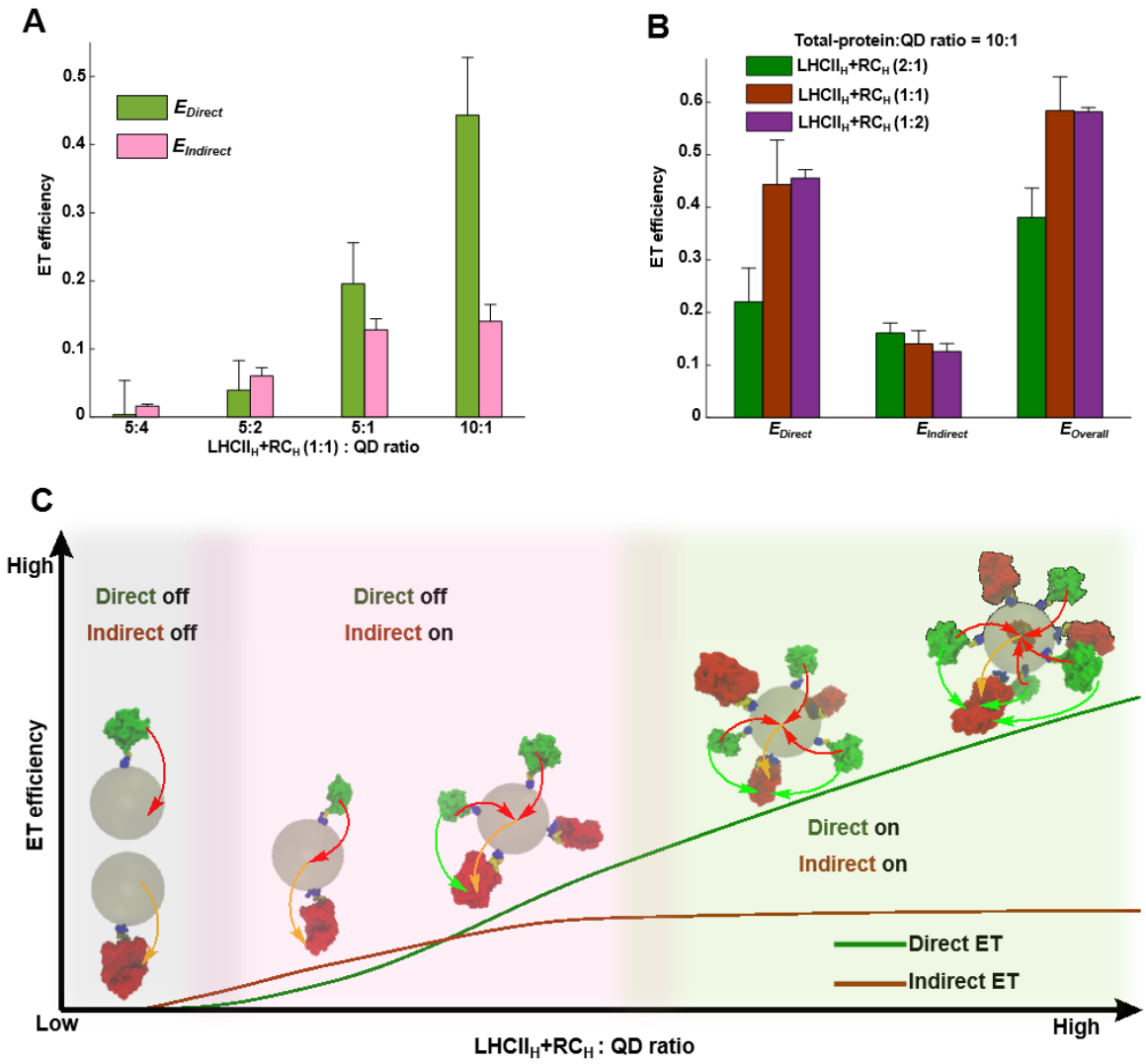


Figure 4.18. Efficiencies and routes of energy transfer in LHCII$_H$/RC$_H$/QD conjugates. (A) Efficiencies of direct and indirect LHCII$_H$→RC$_H$ energy transfer in conjugates formed in mixtures with 1:1 LHCII$_H$:RC$_H$ ratio and four total-protein:QD ratios. (B) Efficiencies of energy transfer in conjugates with three different LHCII$_H$:RC$_H$ ratios at a total-protein:QD = 10:1. $E_{Overall}$ is the sum of E_{Direct} and $E_{Indirect}$. (C) Continuous plots of E_{Direct} and $E_{Indirect}$ against total-protein:QD ratio from 0 (low) to 10 (high) using the trends in panel (A). Conjugates with increasing protein loadings per QD are shown. Energy flow to one RC$_H$ is shown to illustrate the effect of tricomponent conjugate compositions. The grey box indicates the energy transfer-inactive region at very low total-protein:QD ratios. The pink box signifies the switch-on of indirect energy transfer and the start of direct energy transfer as the protein load increases. The green box shows that at high total-protein:QD ratios the direct energy transfer surpasses the indirect route. Green arrows - direct energy transfer (ET); Red arrows - LHCII$_H$ to QD ET; Brown arrows - QD to RC$_H$ ET.

4.3. Discussion

The concept explored in this chapter was to self-assemble, in a predictable way, a nanoscale polychromatic photosystem comprising diverse natural and man-made components as building modules. This was enabled through the use of a His-tag as a universal interface between the two pigment-proteins and the QD assembly hub. The binding affinity determined here for the first $LHCII_H$/QD interaction was similar to that for RC_H/QD conjugation determined in Chapter 3, and spectroscopic analysis showed that the protein content of self-assembled conjugates was determined by the content of the protein mix used to form them up to a stoichiometry about ten protein molecules per conjugate. This predictable self-assembly enabled tuning of the properties of the conjugates in terms of the absolute and relative numbers of $LHCII_H$ light harvesters and RC_H energy traps in a single photosystem despite their being separated by evolution from ancient times.

The interaction of His-tagged LHCII with water-soluble QDs has been previously studied by Werwie and co-workers using the pea Lhcb1.2 polypeptide[151,230]. The recombinant LHCII complexes used in their work also had 11 amino acids removed from the N-terminus but had a His_6 tag placed at either the N- or C-terminus and were induced to trimerize by refolding in the presence of phosphatidylglycerol. Their synthesised CdTe/CdSe/ZnS core-core-shell QDs had a different anionic capping layer (dihydrolipoic acid versus 3-mercaptopropionic acid) and a different surface material (ZnS) to the commercial CdTe QDs used here. Unlike in our study, Werwie and co-workers were able to achieve strong LHCII binding to QDs with a His_6-tag, a difference we attribute to their use of trimeric LHCII, each of which will have three adjacent His_6-tags, and differences in the surface chemistry of the different QDs (Zn^{2+} versus Cd^{2+}). Strong binding of trimeric LHCII to the CdTe/CdSe/ZnS QDs was limited to one per QD, with indications from spectroscopic measurements of a second, weakly bound trimer[151]. In agreement with the findings here, emission from LHCII trimers was quenched by bound QDs indicating FRET between the two.[151,230] The work described in this thesis used LHCII in a monomeric state rather than LHCII trimers in order to be able to bind multiple copies of the $LHCII_H$ to each QD, and to be able to self-assemble an intermingled $LHCII_H$ donor/RC_H acceptor protein shell around the QD hub.

Upon forming $LHCII_H$/RC_H/QD conjugates, the observed photochemistry of each bound RC_H increased as the ratio of $LHCII_H$:RC_H increased (**Fig. 4.9B**) because increased numbers of $LHCII_H$ provided more absorbance cross-section in the red spectrum (650 ± 25 nm, **Fig. S4.5A**) via direct and/or indirect energy transfer. Significant increases in the relative efficiency of direct FRET as the ratio of total-protein:QD increased (**Fig. 4.18A,** green and **Fig. 4.18C,** green trace/green region) are attributed to a reduction in RC_H-$LHCII_H$ separation as the QD

surface became more crowded, bringing RC_H-$LHCII_H$ within the FRET radius. This efficiency rise was sensitive to the $LHCII_H$:RC_H ratio, dropping off when $LHCII_H$ became the minority component and the energy traps outnumbered the light harvesters. Regarding indirect energy transfer, its onset is earlier than direct route (**Fig. 4.18A,** pink and **Fig. 4.18C** brown trace/pink region) caused by a diminishing proportion of FRET-inactive conjugates with exclusively $LHCII_H$-only and RC_H-only protein complements. However this plateaued at a high total-protein:QD ratio because of the competition for the $LHCII_H$ exciton reservoir from direct energy transfer. Summing the direct and indirect components, conjugates with an equal $LHCII_H$:RC_H ratio produced a switch in the overall energy transfer from as low as 0.02 ± 0.05 up to 0.58 ± 0.06 when the total-protein:QD ratio was changed from 5:4 to 10:1 (**Fig. 4.18C**)

The finding that the optimum configuration for this synthetic photosystem was one $LHCII_H$ per RC_H provides an interesting contrast with the efficiency decreases in natural photosystems at increasing antenna size, but at much higher antenna numbers because the effective hierarchy in the organisation of complexes ensuring fast trapping of excitons at the RC[219,231,232]. In another parallel with natural photosynthesis, in addition to their own energy harvesting role the QDs in the tri-component conjugates can be viewed as both a physical and energy bridge between $LHCII_H$ and RC_H, so playing a role similar to that of minor light harvesting components such as CP26 and CP29 in green plant PSII supracomplexes[66,89,182] (see Chapter 1). In the natural system, these minor LHCIIs fill the gap between the large LHCII antenna and the PSII core[233] and contribute to the overall high energy trapping (estimated at ~83 %)[234]. In comparison, the chosen QDs assisted energy flow from Chl-based plant LHCIIs to BChl-based bacterial RCs with a nearly 60% efficiency.

The approach taken here also enabled mimicry of a key aspect of natural photosystems, namely the ability to tune the size of the Chl-antenna to BChl-RC. Examples of the responsiveness of natural photosystems to environmental conditions include acclimation to light intensity by varying the LHCII:PSII ratio and regulating LHCII's distance to PSII or PSI[94], and variation of the size of the peripheral LH2 antenna system per purple bacterial RC in response to light intensity[235]. In the present case, use of a common His-tag interface meant that the average composition of the tri-component conjugates could be controlled by simply varying the relative amounts of $LHCII_H$ and RC_H delivered, providing a means to adapt the response of the conjugates to parameters such as light intensity and spectral composition.

4.4. Conclusions

The research described in this chapter demonstrates the self-assembly of a biohybrid, polychromatic photosystem conjugate comprising the purple bacterial RC, the major plant LHCII antenna and synthetic QDs that displays enhanced solar energy conversion compared to that achievable by the RC alone. In addition to acting as a hub for self-assembly of this bottom-up redesigned photosystem, the QD component fills the energy gap between the plant and bacterial photosystems in a manner analogous to the role played by minor antenna complexes in the multicomponent plant photosystem[89]. The thermodynamic process of self-assembly could be well explained by a competitive binding scheme. Together the direct and indirect FRET produced an efficient energy flow from the LHCII energy harvesters to the RC photochemical traps. Both the overall efficiency and flux down the two routes could be tuned in a predictable manner by simply varying the stoichiometry of $LHCII_H$, RC_H and QD in the mix used for self-assembly. This understanding of the thermodynamics and energy transfer characteristics of RC/LHCII/QD conjugates can be readily transferred to other types of protein/nanocrystal conjugate and will form the basis of future self-assembly of more complex modular bio-conjugates employing a wider range of light harvesting and electron transfer components.

4.5. Supplementary Information

4.5.1. Multiple independent binding model for $LHCII_H$/QD conjugates.

Modelling of the binding of $LHCII_H$ to QDs followed the approach described in Chapter 3 for RC_H/QD conjugates. Average numbers of $LHCII_H$ bound to a QD at different $LHCII_H$:QD ratios ($\bar{\nu}$), calculated from fractions 1-8 in sucrose gradient pulldown experiments (**Fig. S4.11A**), were plotted against the concentration of free $LHCII_H$ [$LHCII_{free}$] calculated from upper fractions 9-10. This plot is shown in **Fig. 4.7B**. These data were then fitted with a model that assumed that the QD provided multiple identical binding sites and that the association of each $LHCII_H$ to a QD was an independent event. The final model is summarised in **Eq. 3.1 in Chapter 3**, where [P_{free}] was the concentration of free $LHCII_H$ ([$LHCII_{free}$]). This was adjusted for the experimental finding that most unbound $LHCII_H$ were damaged ([$LHCII_{damage}$] for [P_{damage}] in **Eq. 3.1**) as indicated by a change of the absorbance spectrum of this fraction which is a signature of disruption of structure[236] (**Fig. S4.11B**). The terms deduced from the fit were the microscopic thermodynamic association constant (k_{micro}), the maximum number of $LHCII_H$ binding to each QD ($\bar{n}$) and amount of damaged $LHCII_H$ expressed as a percentage of the total $LHCII_H$ (i.e. $LHCII_{damaged}$ = [$LHCII_{damage}$]/[$LHCII_{total}$]). The best fit produced $\bar{n}$ = 17, k_{micro} = 7.19 μM^{-1} and $LHCII_{damaged}$ = 0.057.

Parameters k_{micro} and $\bar{n}$ were then used to determine macroscopic dissociation constants (K_d) at permitted valencies (i) from **Eq. 3.2**, producing a value of 8.2 nM for i = 1. Conjugate assembly and disassembly were addressed through the model depicted in **Fig. 4.7A** and the reaction scheme shown in **Eq. 3.3**, where P is $LHCII$, and then $LHCII_{His}$ is the total $LHCII_H$ population adjusted for damaged $LHCII_H$ unable to bind and $k_{f,i}$ and $k_{r,i}$ are the macroscopic kinetic constants for binding and dissociation, respectively, at permitted valencies (i) between 0 and 17. As illustrated in **Fig. 4.7A,** parameters $k_{f,i}$ and $k_{r,i}$ were determined from the corresponding microscopic kinetic constant according to **Eq. 3.6**, **3.7**. Deterministic simulations (**Eq. 3.4, 3.5**) based on the reaction scheme (**Eq. 3.3**) were carried out using the defined kinetic constants until equilibria were reached (**Fig. S4.12**). Distributions of the binding stoichiometry were extracted after the simulation had relaxed to equilibrium (**Fig. S4.13**).

4.6. Supplementary figures

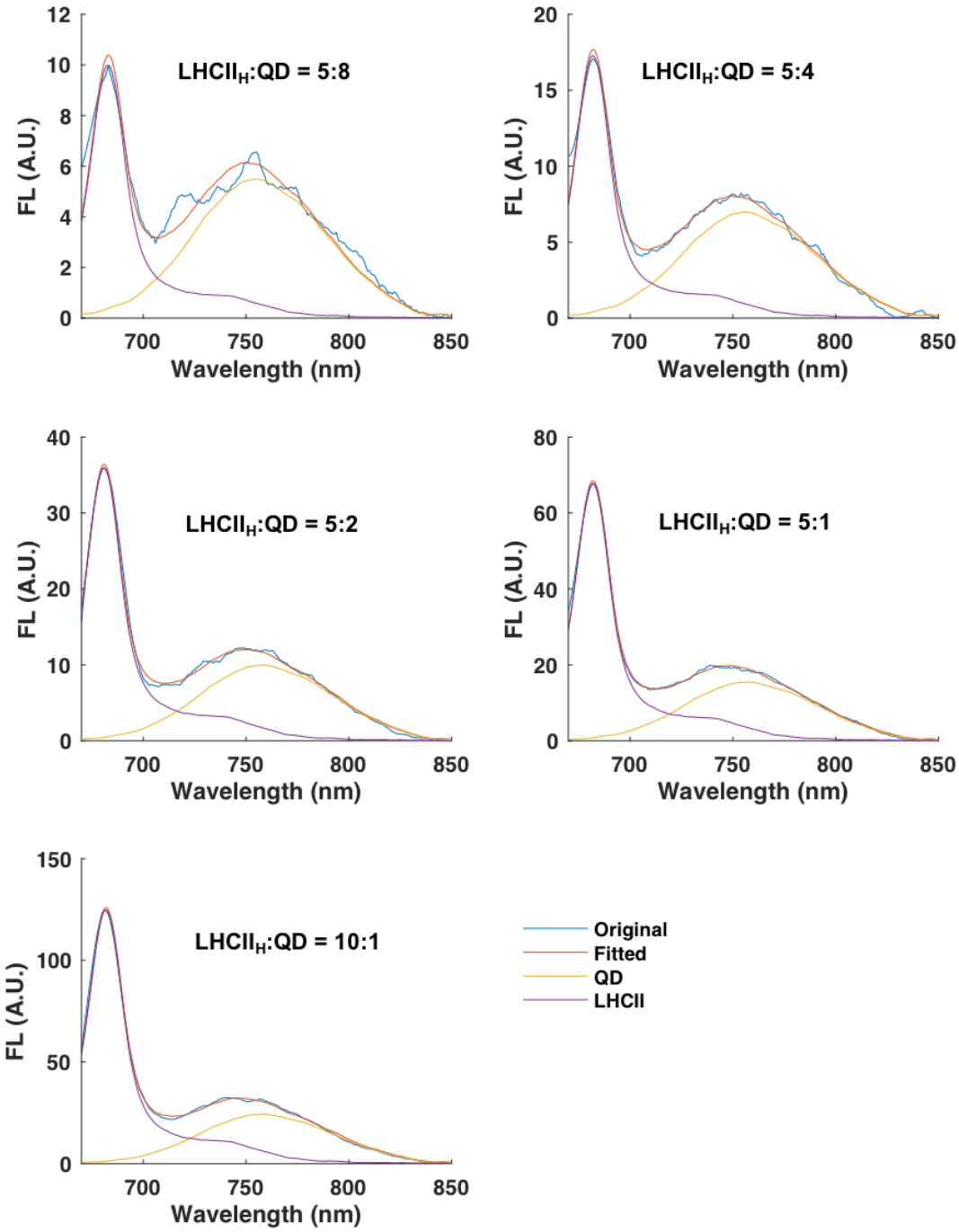


Figure S4.1. Deconvolution of emission spectra from LHCII$_H$/QD conjugates. Raw fluorescence spectra were baseline corrected, smoothed and deconvoluted into LHCII$_H$ and QD components.

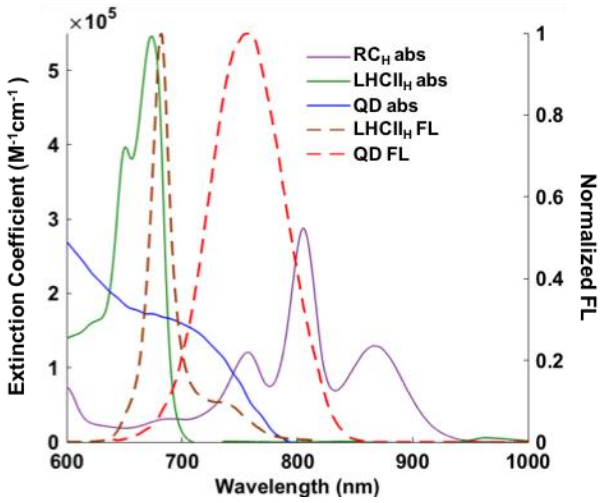


Figure S4.2. Determination of spectral overlap. Molar absorption spectra of $LHCII_H$, RC_H and QDs, and corresponding normalized emission spectra.

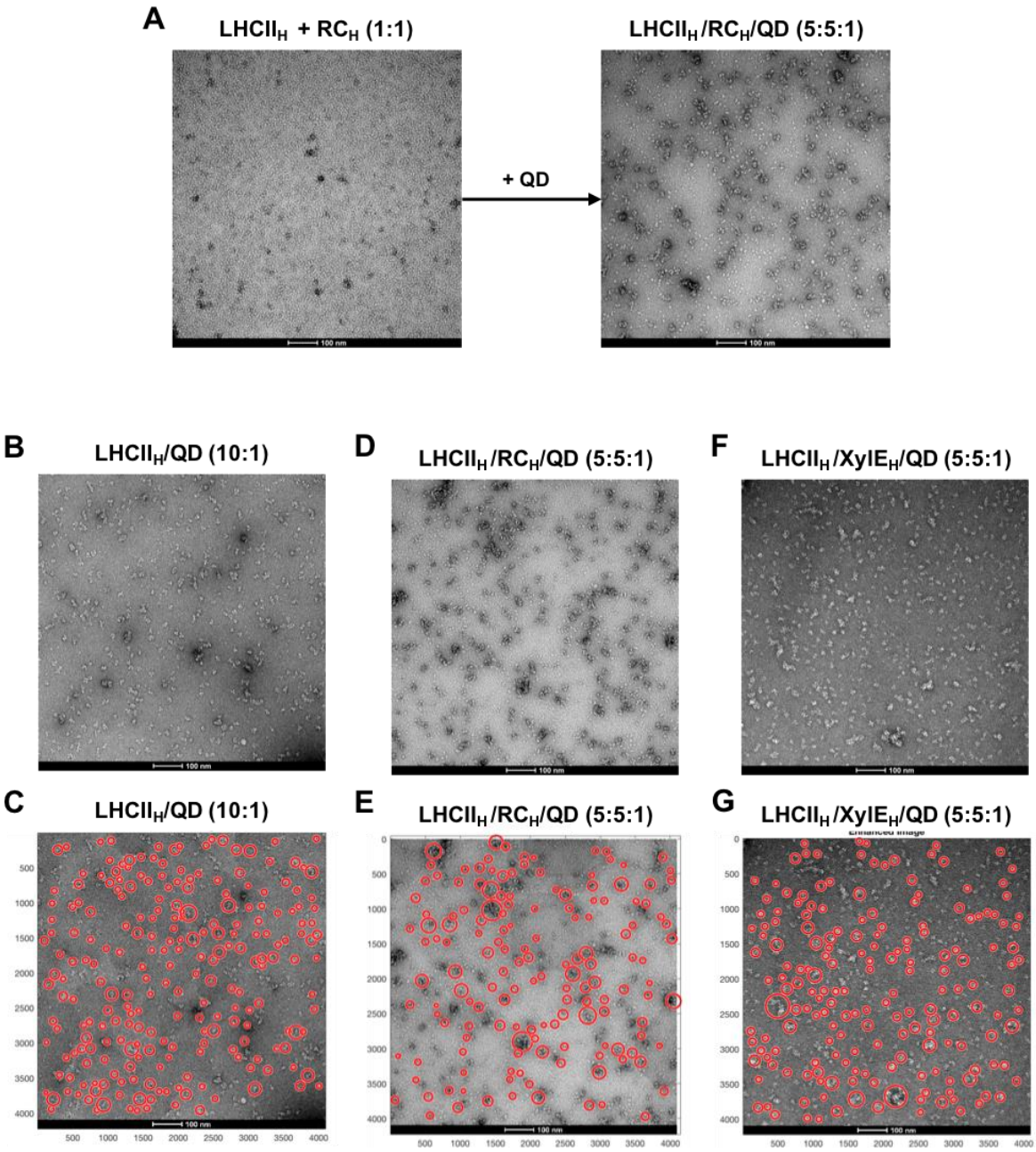


Figure S4.3. TEM of conjugates. (A) Change in morphology of a 1:1 mix of LHCII$_H$ and RC$_H$ proteins induced by adding 0.1 molar equivalent of QDs. The resulting larger conjugates presented as monodispersed objects with a grossly round shape and a diameter that was consistent with a shell of pigment-proteins surrounding a central QD. (B) LHCII$_H$/QD conjugates formed from a 10:1 molar mix and (C) detected objects highlighted with red circles. (D) Equivalent data for LHCII$_H$/ RC$_H$/QD tricomponent conjugates from a 5:5:1 molar mix and (E) detected objects highlighted with red circles. (F) Equivalent data for LHCII$_H$/XylE$_H$/QD tri-component conjugates from a 5:5:1 molar mix and (G) detected objects highlighted with red circles.

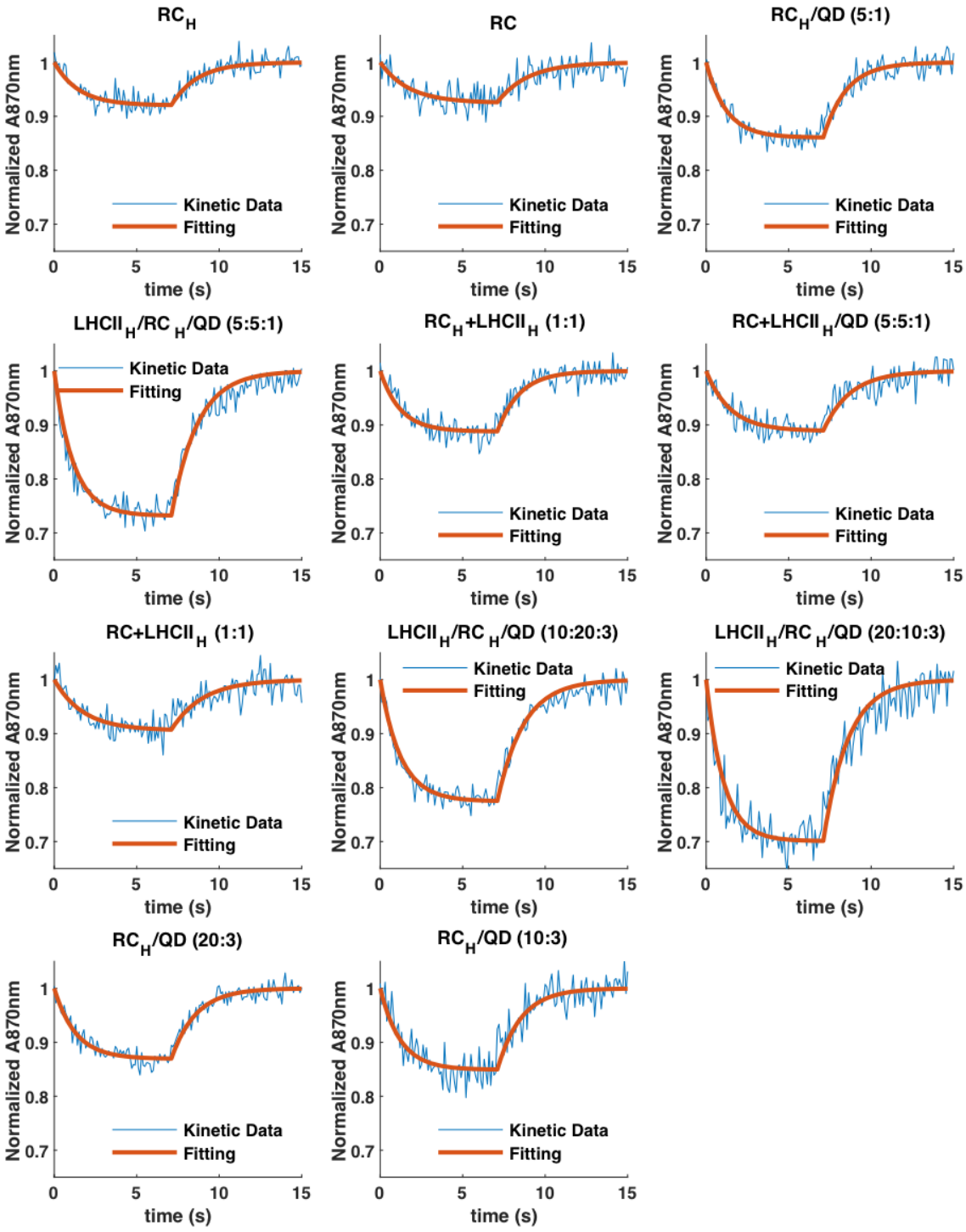


Figure S4.4. RC P870 photobleaching, dark recovery and fitted kinetics. Each panel shows averaged kinetic traces and single exponential fits to the dark and light phases. Excitation was through a bandpass filter centred at 651 nm for 7 seconds followed by dark recovery. The composition of each sample is indicated above each panel. Ratios are for the molar mix used to form the conjugate.

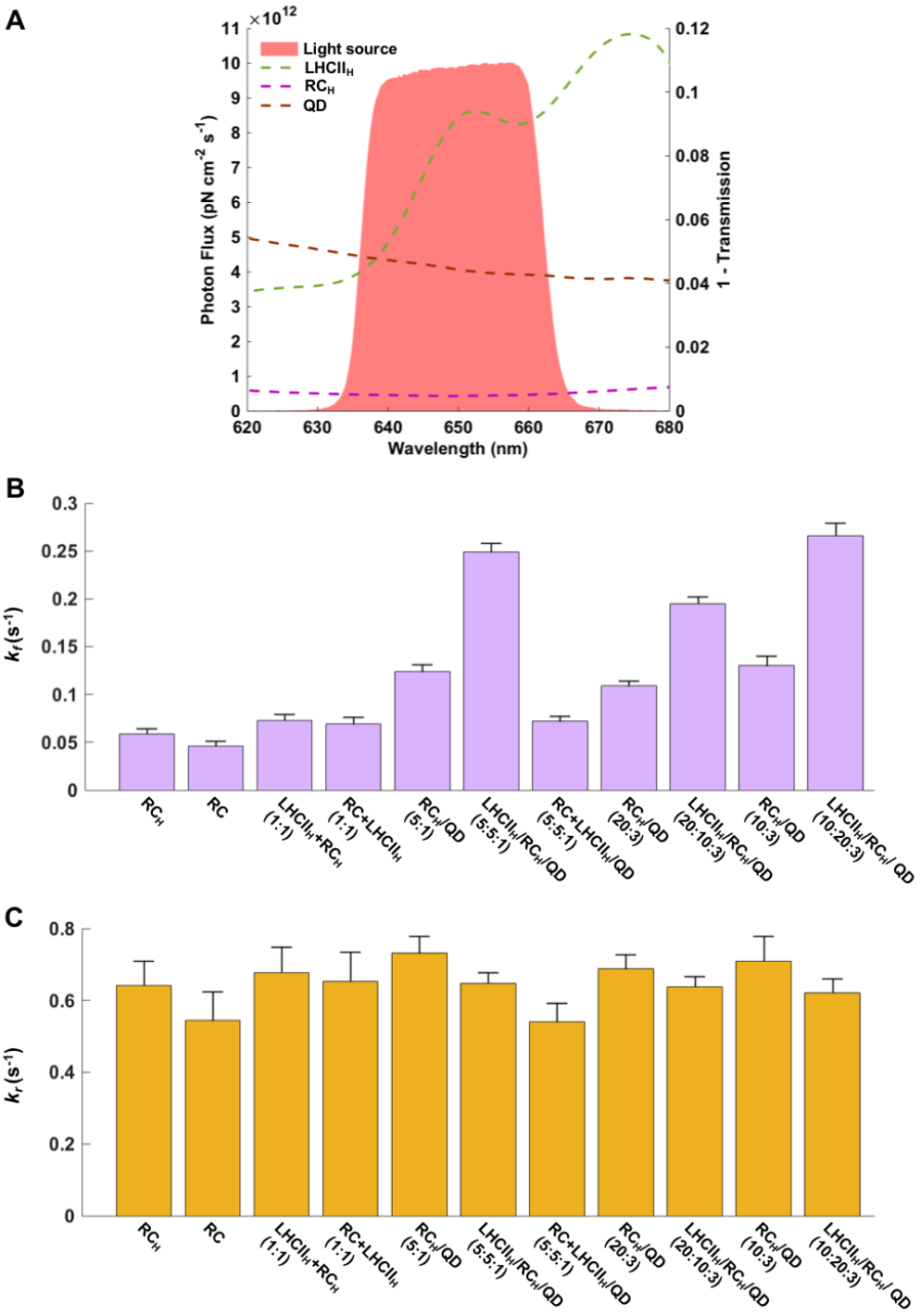


Figure S4.5. Kinetics of RC P870 photobleaching and dark recovery. (A) Illumination light profile in units of photon flux overlaid with the absorbance spectra of RC_H, $LHCII_H$ and QD all at 100 nM. The light source primarily excited the Chl *b* pigments of $LHCII_H$ and the QDs. (B) Rate of P870 photobleaching (k_f) in response to excitation. (C) Rate of dark recovery (k_r) of P870 photobleaching.

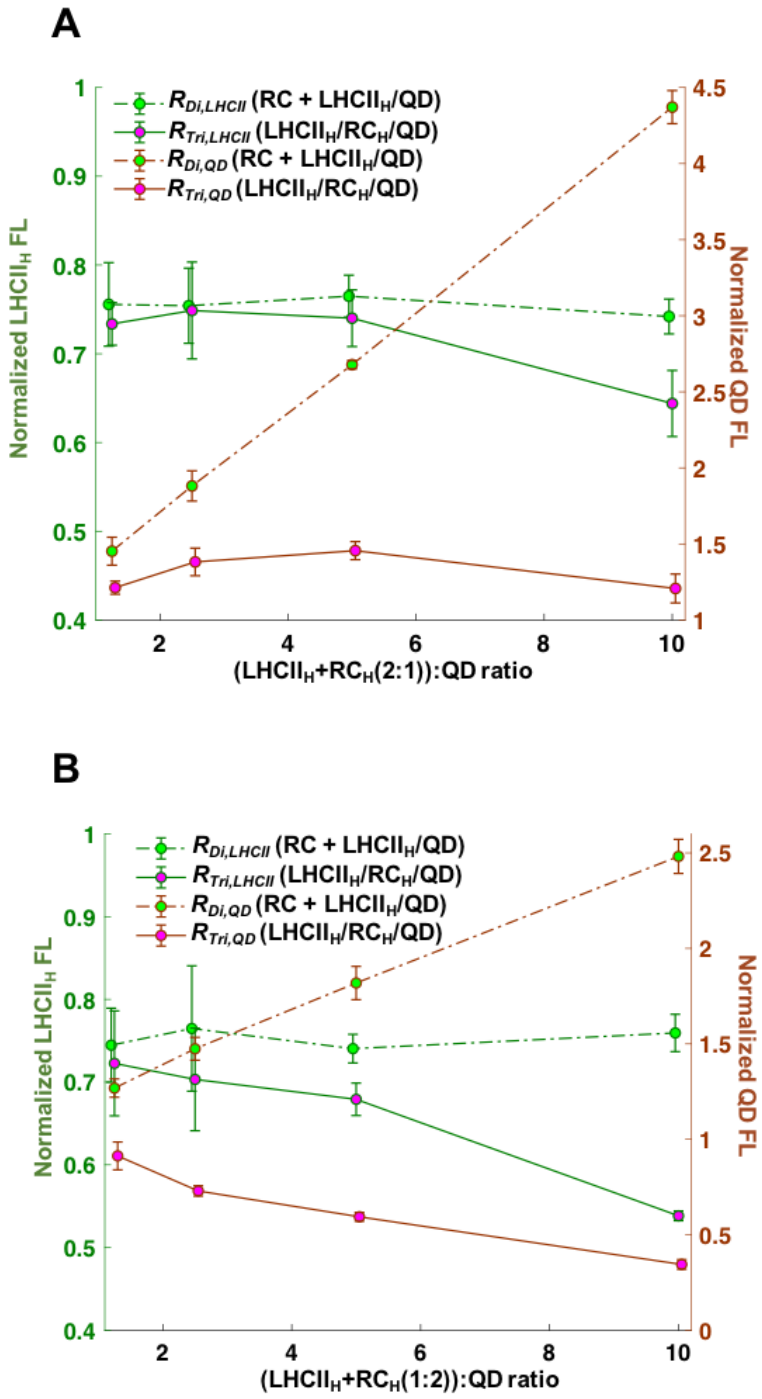


Figure S4.6. Energy transfer in tri-component conjugates. Normalised $LHCII_H$ or QD emission as a function of the ratio of total-protein to QD for $LHCII_H$/RC_H/QD conjugates or a composition-matched mixture of non-His-tagged RC and $LHCII_H$/QD conjugate. After deconvolution, $LHCII_H$ or QD emission was normalised to the emission from an equivalent concentration of $LHCII_H$-only or QD-only single component samples. Data are shown for a 2:1 $LHCII_H$:$RC_{(H)}$ mix (A) and a 1:2 $LHCII_H$: $RC_{(H)}$ mix (B). Data for a 1:1 $LHCII_H$:$RC_{(H)}$ mix are shown in **Figure 4.4D**.

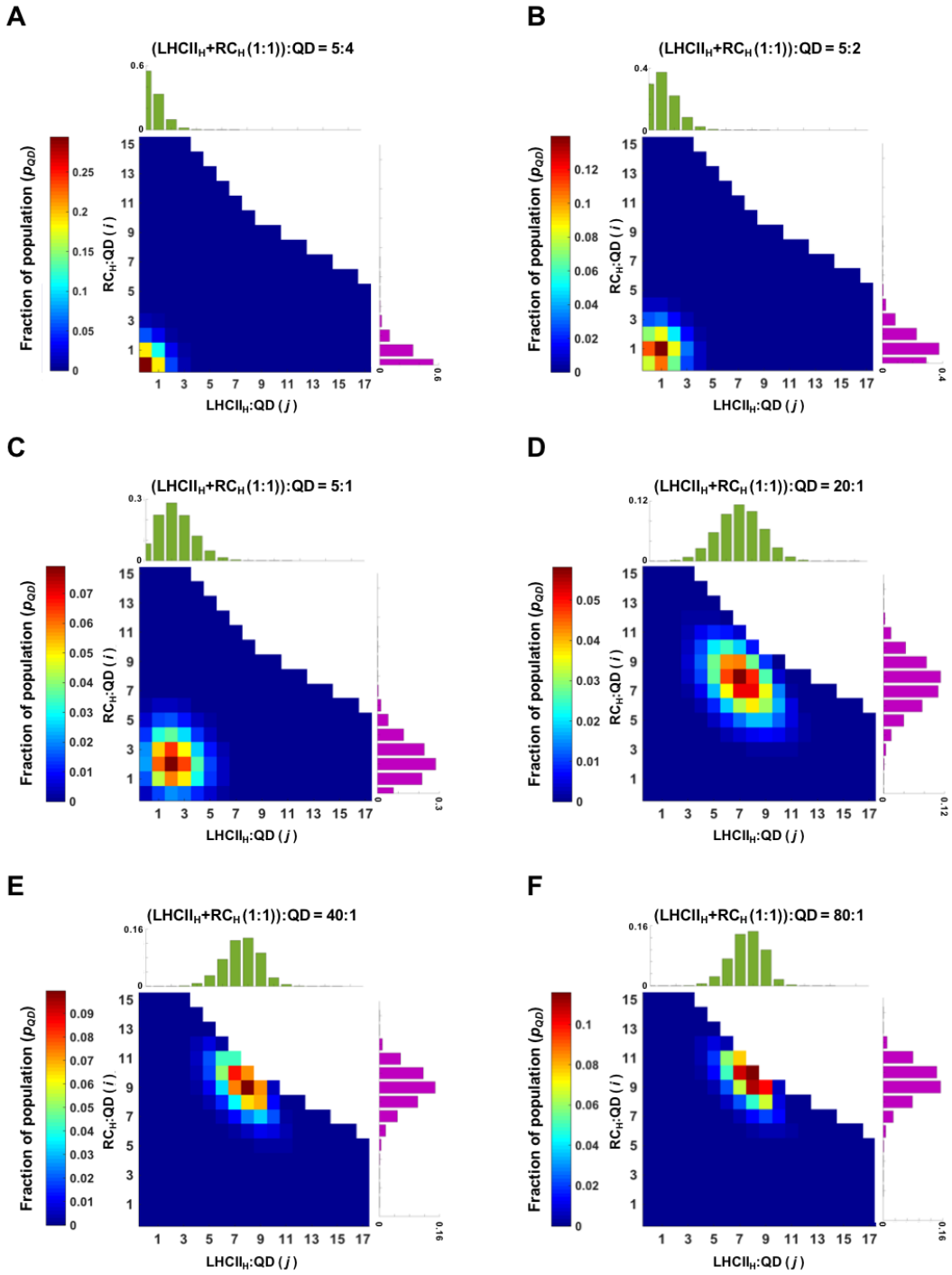


Figure S4.7. Computed population heterogeneities for LHCII$_H$/RC$_H$/QD conjugates. (A-F) Population heterogeneities for conjugates formed from six different mixes in which LHCII$_H$:RC$_H$ was fixed at 2.5μM (1:1) and total-protein:QD varied between and 1.25:1 and 80:1. Total-protein:QD ratios in the initial mix are indicated at the top of each panel. Individual distribution probabilities of LHCII$_H$ per QD and RC$_H$ per QD are shown with green bars and purple bars.

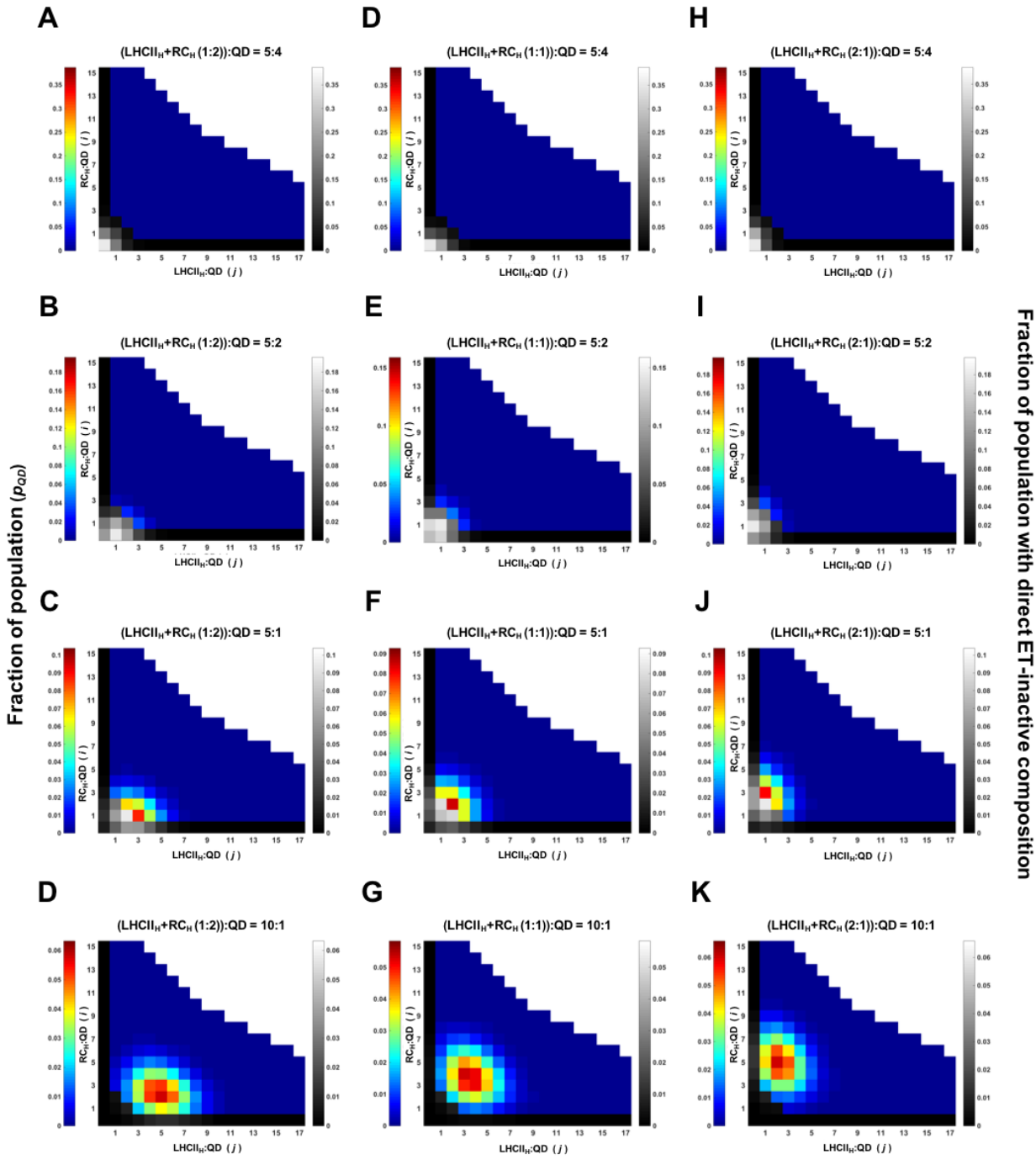


Figure S4.8. Estimations of the sub-population of LHCII$_H$/RC$_H$/QD conjugates inactive in FRET. Data are computed for conjugates in mixes with LHCII$_H$:RC$_H$ ratios of 1:2 (left), 1:1 (centre) and 2:1 (right) and total-protein:QD ratios of 1.25 (top row) to 10:1 (bottom row). Conjugate compositions that are considered to be active in FRET are shown in colour and those that are inactive are shown in greyscale. The latter include all conjugates within the population where the number of RC$_H$ or LHCII$_H$ bound to a QD is zero (left column and bottom row within each plot). The remainder of the inactive zone is in the bottom-left of each plot, based on the observation of no energy transfer at a total-protein:QD ratio equal to 1.25. The

population distribution begins to migrate out of the energy transfer inactive zone at a total-protein:QD ratio of 2.5:1 (second row) and has migrated out of this zone at a total-protein:QD ratio of 10:1 (bottom row). Moving the $LHCII_H$:RC_H ratio move away from one (left and right columns) resulted in a bigger subpopulation incapable of energy transfer, a phenomenon that persisted up to a total-protein:QD ratio of 10:1 (D, K compared to G). This may have contributed to the observed lower energy transfer efficiencies at $LHCII_H$:RC_H ratios other than one.

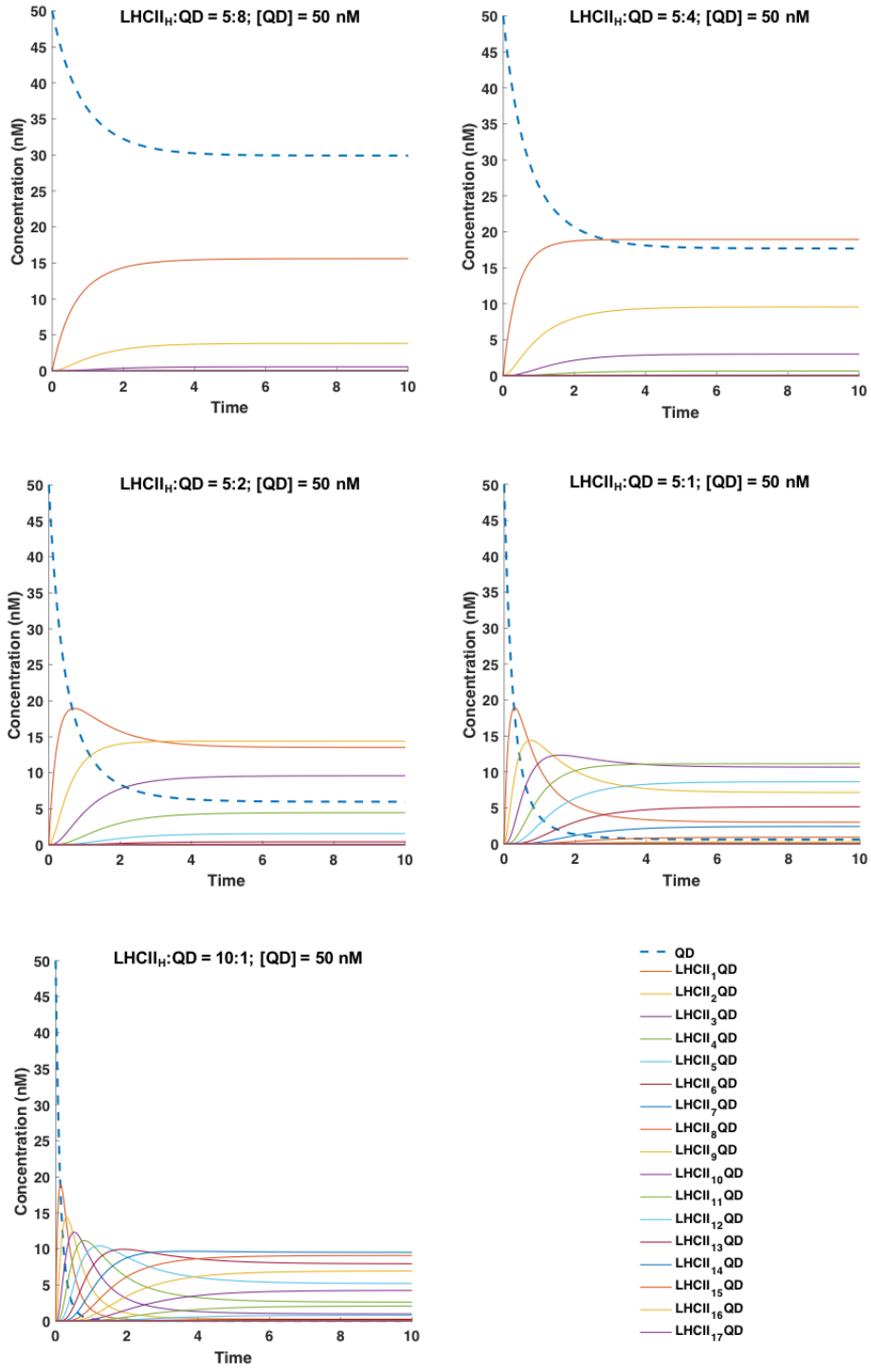


Figure S4.9. Simulation of heterogeneity of LHCII$_H$/QD conjugate composition. Simulations over time of the concentrations all possible LHCII$_{H(i)}$/QD conjugates (*i = 0, 1, …, 17*) formed in mixes with different overall LHCII$_H$:QD ratios. Initial conditions are defined on top of each panel. The concentration of free QD was shown in dashes and those of conjugates in solid lines.

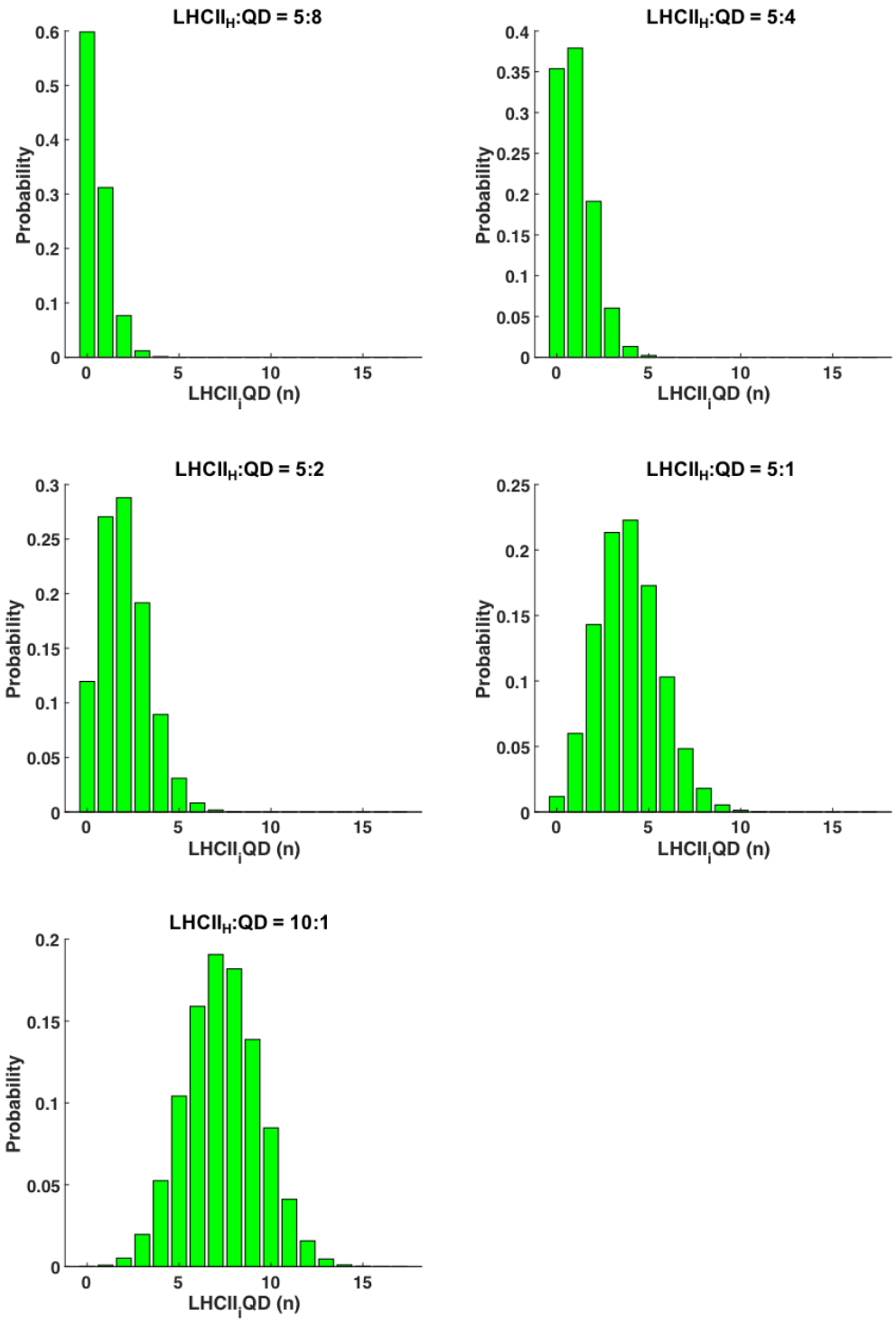


Figure S4.10. Final fraction distribution of LHCII$_H$/QD conjugate composition in simulations. Final probabilities of the formation of LHCII$_{H(i)}$/QD conjugates (*i = 0, 1, …, 17*), expressed as the fraction of the total population, derived from the simulations shown in **Fig. 4.9**. The starting LHCII$_H$:QD ratios are stated above each panel.

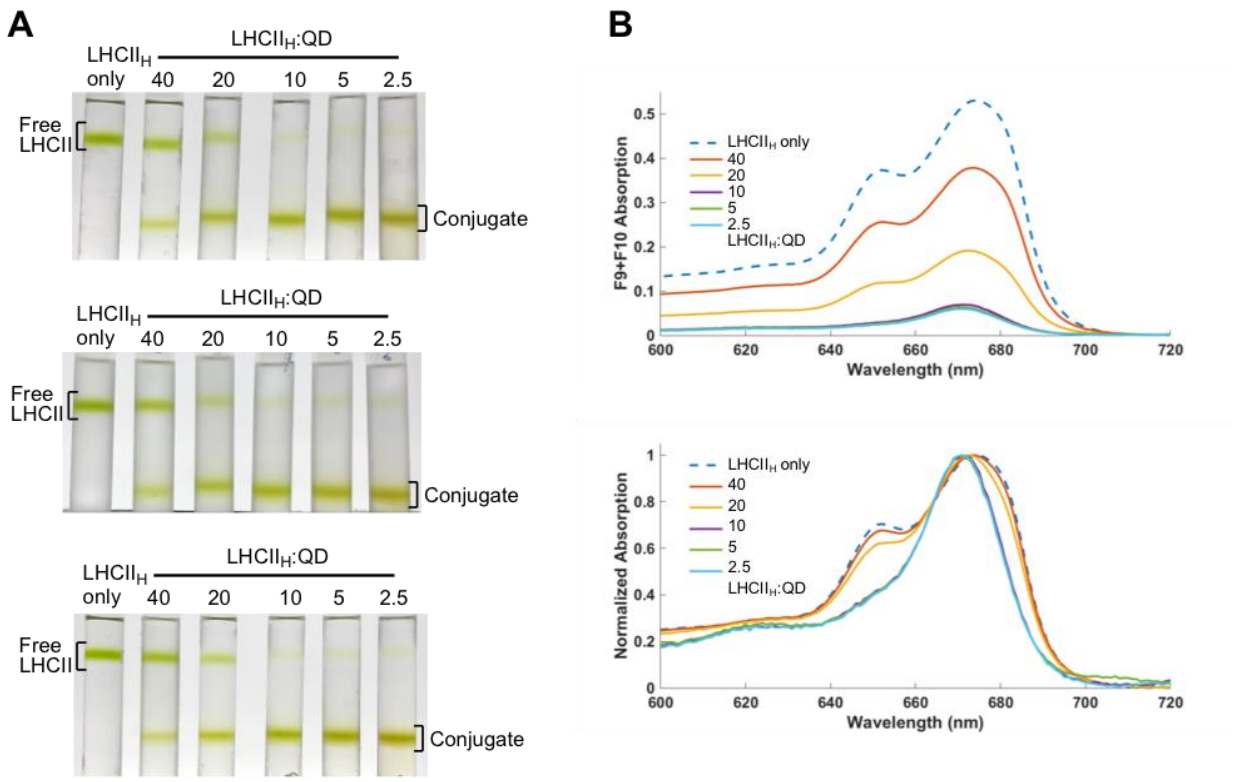


Figure S4.11. Sucrose gradient ultracentrifugation of LHCII$_H$/QD conjugates. (A) Three repeats of "sucrose pull-down" experiments to separate free LHCII$_H$ from conjugates (positions highlighted) on 0%/25%/65% sucrose gradients. The concentration of LHCII$_H$ in all gradients was fixed at 2.5 µM and the concentration of QDs varied between 1 µM and 62.5 nM to give the desired LHCII$_H$:QD ratio. (B) Absorbance spectra of pooled upper fractions 9/10 containing unbound LHCII$_H$ (top) and the same spectra normalized to the Chl *a* Q_y band (bottom). The spectrum of the LHCII$_H$ control is plotted in dashes. The change in line shape on decreasing LHCII$_H$:QD ratio was due to the upper fraction becoming increasingly dominated by damaged LHCII$_H$ incapable of binding to QDs. This fraction was estimated to be 5.7 % of the total LHCII$_H$ population (see main text).

Chapter 5

Polychromatic Solar Energy Conversion in Pigment-Protein Chimeras that Unite the Two Kingdoms of Bacterio-Chlorophyll Photosynthesis

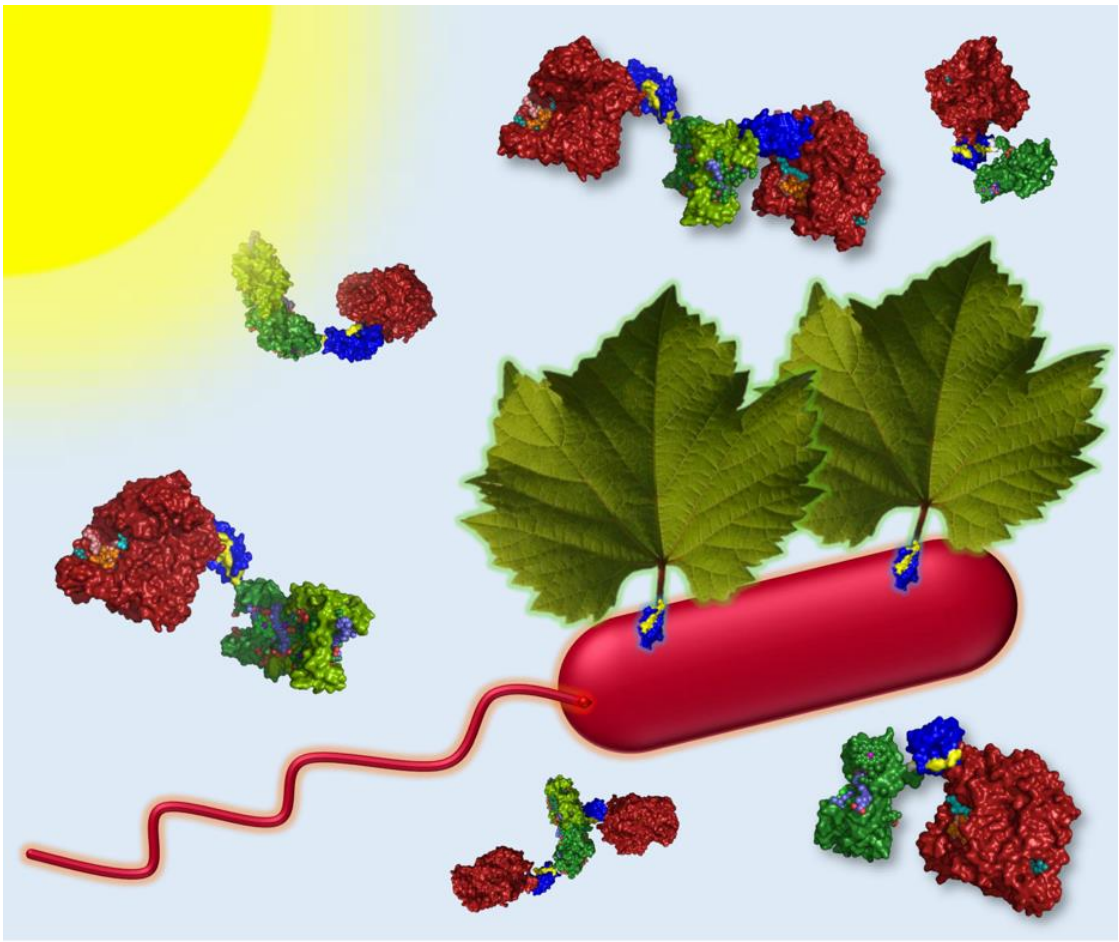
A manuscript describing the results presented in the chapter has been submitted for publication and was written by me:

Liu, J., Friebe, V., Frese, R.N. and Jones, M.R. (2019) Polychromatic solar energy conversion in pigment-protein chimeras that unite the two kingdoms of bacterio-chlorophyll photosynthesis.

The experiments and analysis described in this chapter were performed by me with the exception of measurement of photocurrents and EQE action spectra which were performed jointly with Dr. Vincent Friebe at the Free University Amsterdam during a research visit to the laboratory of Dr. Raoul Frese.

Abstract

Natural photosynthesis can be divided between the Chl-containing plants, algae and cyanobacteria that make up the oxygenic phototrophs and a diversity of BChl-containing bacteria that make up the anoxygenic phototrophs. Photosynthetic light harvesting and RC proteins from both groups of organisms have been exploited in a wide range of biohybrid devices for solar energy conversion, solar fuel synthesis and a variety of sensing technologies, but the energy harvesting abilities of these devices are limited by each protein's individual palette of (bacterio)chlorophyll, carotenoid and bilin pigments. In this work we demonstrate a range of genetically-encoded, self-assembling photosystems in which recombinant plant LHCs are covalently locked with RCs from a purple photosynthetic bacterium, producing macromolecular chimeras that display mechanisms of polychromatic solar energy harvesting and conversion not present in natural systems. Our findings illustrate the power of a synthetic biology approach in which bottom-up construction of a novel photosystem using naturally disparate but mechanistically complementary components is achieved in a predictable fashion through the genetic encoding of an adaptable, plug-and-play covalent interface.



5.1. Introduction

Our everyday experience of photosynthesis is dominated by blue/red-absorbing Chl *a* and Chl *b* found in the plants, algae and cyanobacteria that surround us. Less obvious are the anoxygenic phototrophs, bacteria that have one or more variants of BChl as their principal photosynthetic pigment. Although these bacteria are less obvious in our environment, oxygen-tolerant species are widespread in oceanic surface waters where they make a sizeable contribution to global solar energy conversion[237].

As outlined in Chapter 1, improving the performance of photosynthesis and finding new ways to exploit natural solar energy conversion have become important research topics[125,127], and there is growing interest in the use of photosynthetic proteins as environmentally-benign components in biohybrid devices for solar energy conversion[91,105,107,109,222,238]. Photoexcitation of a RC in such a device triggers intra-protein charge separation, producing a potential difference between opposite "poles" of the protein that drives subsequent electron transfer to create a photocurrent and photovoltage. In addition to solar energy conversion *per se*, proposed applications of photoprotein devices include biosensing, light/UV sensing, touch sensing and solar fuel synthesis[91,105,107,109,115,120,222,238]. Photosynthetic proteins are attractive as device components because they are environmentally sustainable and benign, they achieve solar energy conversion with a very high quantum efficiency and they can be adapted to purpose through protein engineering. However, a limitation is their selective use of available solar energy[125,127], a consequence of their particular palette of light harvesting pigments. This can be evidenced in devices through the recording of action spectra of external quantum efficiency (EQE – the number of charges transferred per incident photon), which exhibit peaks and troughs that correspond to the absorbance spectra of the particular light harvesting pigments that are coupled to charge separation in the device[109,134,139,140,193,194].

As described in Chapters 1 and 4, a striking observation is the complementary nature of the absorbance spectra of Chl and BChl photosystems (**Fig. 1.2B**). In Chapter 4 this was exploited by combining plant LHCs and purple bacterial RCs into tri-component conjugates with CdTe QDs, producing a semi-synthetic photosystem that showed efficient transfer of energy from LHCs to RCs. Although the CdTe QDs were effective both as a tight-binding assembly hub and an energy bridge, the drawbacks of their use are that cadmium is toxic, the binding interactions that hold the conjugate together are non-covalent, the binding could interfere with access of electron donor and acceptors (including electrodes) to the oxidising and reducing poles of the RC, and commercial CdTe QDs are very expensive. In this chapter,

an alternative all-biological mechanism for bringing plant LHCs and purple bacterial RCs together into a single photosystem is explored.

Following nature's lead, this chapter describes the use of genetic encoding to achieve the self-assembly of novel photoprotein "chimeras" that display polychromatic solar energy harvesting and conversion. The components used were the *Rba. sphaeroides* RC[38,51] and the LHCII[64–66,239] and LHCI[44,73–76] pigment-proteins from *A. thaliana* (**Fig. 5.1A-C**). Using an approach inspired by synthetic biology, highly specific and programmable self-assembly is achieved through adaptation of these components with the constituents of a two-component protein interface domain that covalently locks together two or three photosynthetic membrane proteins which have no natural affinity for one another. This interface domain was provided by the SpyCatcher/SpyTag system (**Fig. 5.1D**) described in Chapter 1.

The resulting macromolecular, adaptable chimeric photosystems assemble with high yield, have highly defined and stable compositions, and display novel mechanisms of solar energy conversion across the near-UV, visible and near-IR both in solution and on an electrode surface.

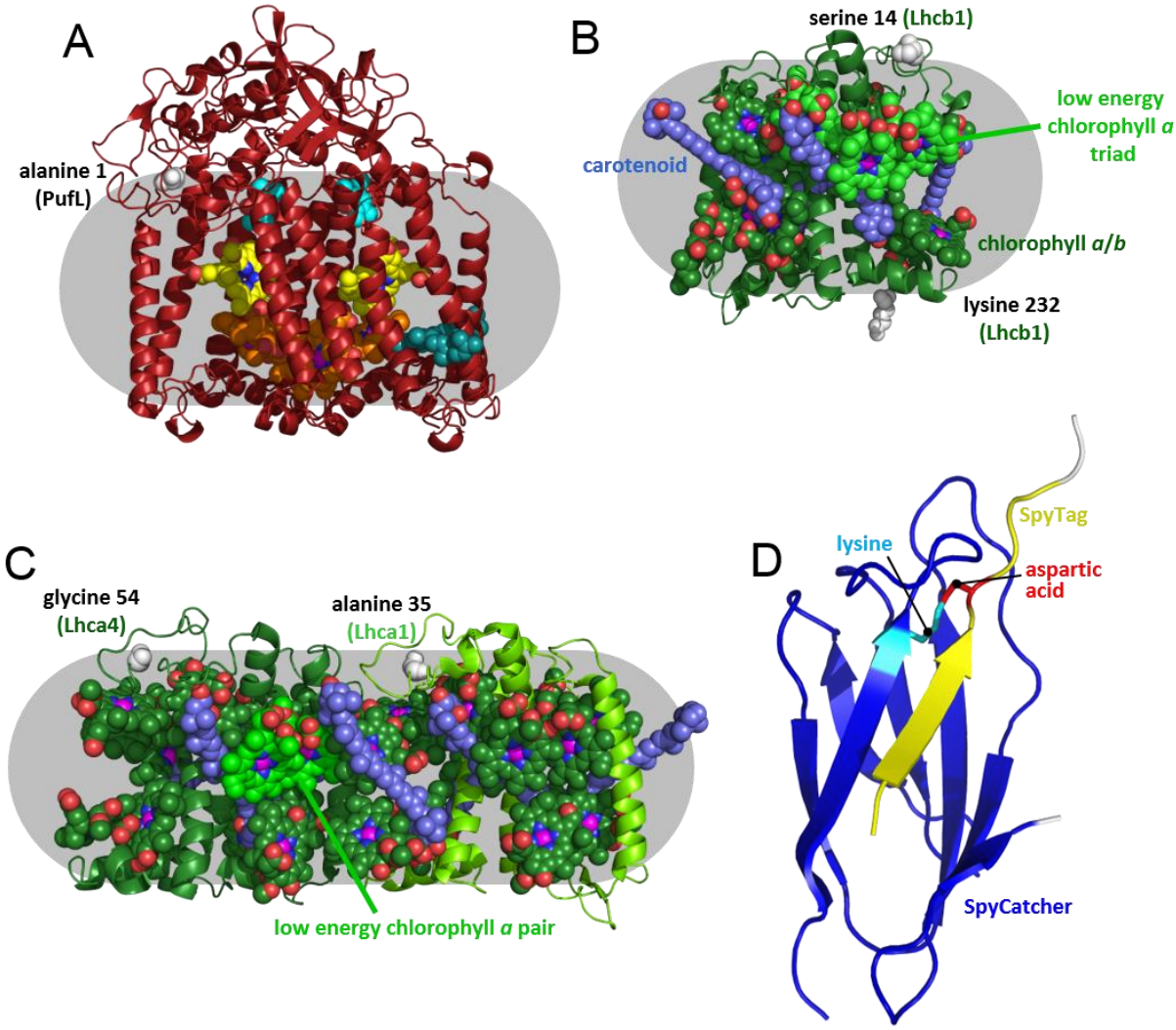


Figure 5.1. Component structures. (A) The *Rba. sphaeroides* RC comprises three polypeptides (maroon ribbons) that scaffold its cofactors. Grey shading approximates the dimensions of a detergent micelle. Non-native polypeptides were added before alanine 1 of the PufL subunit (white spheres, top left). (B) LHCII comprises a single Lhcb1 polypeptide (green ribbon) that scaffolds 8 Chl *a*, 6 Chl *b* (green carbons) and 4 carotenoids (slate carbons). The carbons of three interacting Chl *a* that from a low energy cluster are highlighted in light green. Non-native polypeptides were added before serine 14 or after lysine 232 (white spheres). (D) LHCI complexes formed from a heterodimer of Lhca1 and Lhca4 (light/dark green ribbons) that scaffold 36 cofactors (23 Chl *a*, 6 Chl *b* and 7 carotenoids). The carbons of two interacting Chl *a* that from a low energy cluster are highlighted in light green. Non-native polypeptides were added before alanine 35 of Lhca1 or glycine 54 of Lhca4 (white spheres) which are the N-terminal residues in the resolved X-ray structure. (D) SpyTag (yellow) binds to SpyCatcher (blue) and an isopeptide bond is formed between lysine (cyan) and aspartic acid (red) side-chains. Residues removed from the C-terminus of SpyTag are coloured white (top right). Sources of the structures are listed in **Figs. 5.7** and **5.9**

5.2. Results

5.2.1. Methodologies

As with work described in previous chapters, genetically-modified RCs were expressed in a LHC-deficient strain of *Rba. sphaeroides* and purified by nickel affinity chromatography (Sections 2.5 and 2.6). Genetically-modified LHCII and LHCI pigment proteins were refolded *in vitro* using purified pigments and apoproteins that had been expressed in *E. coli* (Sections 2.4 and 2.6).

5.2.2. Solar energy conversion by unadapted photosystem components

The first step was to look at the extent to which unadapted plant LHCIIs can pass harvested energy to unadapted purple bacterial RCs in dilute solution. As explained in previous chapters, on receipt of excitation energy, photochemical charge separation in the *Rba. sphaeroides* RC is a rapid four-step process that produces a metastable oxidised primary electron donor ($P870^+$) and reduced acceptor ubiquinone (Q_B^-). Energy transfer can therefore be detected as a quenching of LHC emission accompanied by an enhancement of P870 oxidation. Although unadapted WT *Rba. sphaeroides* RCs and *A. thaliana* LHCIIs have overlapping absorbance and emission spectra between 600 nm and 800 nm (**Fig. 5.2A**; LHCII is named dLHCII – see below) no appreciable energy transfer was observed when they were mixed in solution because they have no capacity for binding to one another. The addition of purified WT RCs did not significantly reduce emission from LHCII (**Fig. 5.2B**) and photo-oxidative bleaching of this RC's P870 primary electron donor absorbance band in response to 650 nm excitation was not significantly enhanced by the addition of LHCII (**Fig. 5.2C**) which absorbs strongly at this wavelength (**Fig. 5.2A**).

In comparison to LHCII, the spectral overlap (*J*) between LHC emission and RC absorbance is ~80 % larger in the case of LHCI (**Fig. 5.3A**) which contains a pair of "red-form" Chl *a* that possess a charge-transfer state that mixes with the low energy exciton state[76]. Spectral overlaps are quantified later in this chapter, as are the quantum yields of all LHCs studied. Although the addition of WT RCs did bring about a decrease in LHCI emission (**Fig. 5.3B**) there was no associated significant increase in RC P870 photobleaching in the presence of LHCI (**Fig. 5.3C**), leading to the conclusion that the observed emission quenching was not due to energy transfer. Protein concentrations used for the fluorescence measurements were too low (max absorbance < 0.07) for this LHCI quenching to be attributable to reabsorption by RCs, and equivalent quenching was not seen with LHCII at a similar concentration (**Fig. 5.2B**). As it is known that the emission quantum yield of LHCI is much more sensitive to its *in vitro*

environment than is the case for LHCII[74], the observed drop in emission on adding RCs is attributed to a change in the intrinsic quantum yield of LHCI.

To establish the principle that plant LHCs can pass energy to bacterial RCs when brought sufficiently close together, mixtures of unadapted LHC and RC proteins were adhered to a nanostructured silver cathode and their capacity for generating photocurrents examined. The methodology used for these measurements is described in Section 2.18. In the photoelectrochemical system used (**Fig. 5.4A**) horse heart cyt *c* is used to "wire" charge separation in the RC to the electrochemically-roughened silver cathode, and ubiquinone-0 (Q_0) shuttles electrons to the counter electrode[91,140].

Nanostructured silver electrodes drop-cast with purified WT RCs produced a photocurrent in response to RC-specific 870 nm light and a weaker current in response to 680 nm excitation where RC absorbance is very low (**Fig. 5.4B**, magenta and blue). To look at the source of the photocurrent, external quantum efficiency (EQE) action spectra were recorded. EQE is the ratio of the number of electrons in the external circuit produced by an incident photon of a given wavelength and is calculated by essentially scaling the photocurrent density by the illumination intensity at each wavelength. The result is an action spectrum that reveals the wavelengths of light being used to drive the measured photocurrent. The RQE spectrum recorded for an electrode with adhered WT RCs showed good correspondence with the RC absorbance spectrum, confirming that the photocurrent was attributable to light capture by the pigments of the RC (**Fig. 5.4D**, black versus red).

As expected, an electrode fabricated with purified LHCII (or LHCI) failed to show any photocurrent response during excitation of the main LHCII absorbance band at 680 nm (**Fig. 5.4B**, green) or at any other wavelength. However, when LHCII and WT RCs were mixed on the electrode surface the photocurrent produced by 680 nm excitation was larger relative to that obtained with 870 nm excitation (**Fig. 5.4C**) than was the case for RCs alone (**Fig. 5.4B**). Variations in the absolute level are likely attributable to variations in loading on each electrode, which was not quantified.

EQE spectra were also recorded for electrodes with a mix of WT RCs and LHCII or a mix of WT RCs and LHCI. In both cases in addition to the expected RC bands the EQE spectra contained a component between 620 and 700 nm that corresponded to the low energy absorbance band of LHCII or LHCI (**Fig. 5.4E**, dark/light green). A contribution from the high energy Soret absorbance band of LHCII or LHCI was also evident (**Fig. S5.1B,C**) when EQE spectra were recorded using a blue-enhanced xenon light source (**Fig. S5.1A**).

The line shapes of the EQE spectra demonstrated that photoexcited LHCII or LHCI Chls can pass energy to the BChls of a purple bacterial RC, producing charge separation and a photocurrent response. In this case proximity-dependent energy transfer was realised by bringing the two types of protein sufficiently close to one another on the surface of a biophotoelectrode.

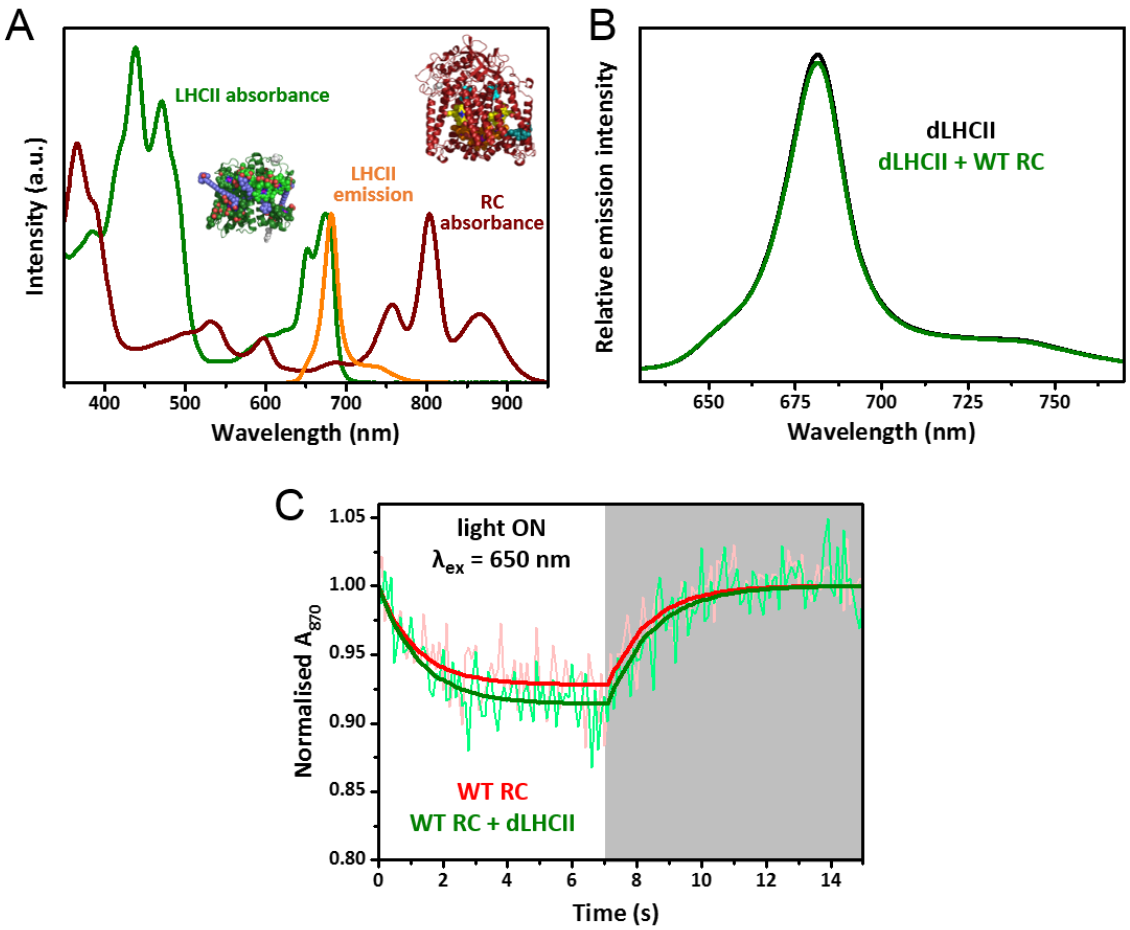


Figure 5.2. Mixing unadapted LHCII and RCs in solution. (A) LHCII harvests solar energy in regions where absorbance by *Rba. sphaeroides* RCs is weak, and its emission spectrum overlaps the absorbance spectrum of the RC between 600 nm and 800 nm. (B) LHCII emission in the absence and presence of WT RCs (excitation at 475 nm). (C) Data and fits for photobleaching and dark recovery of P870 absorbance for the WT RC in the absence and presence of LHCII. Kinetic constants from the fits are summarised in supplementary **Table S5.1**. Excitation was for 7s through a ~50 nm wide bandpass filter centred at 650 nm. In all panels the variant of recombinant LHCII used had 12 amino acids removed from the N-terminus of Lhcb1 and a His-tag at the C-terminus (named dLHCII – see text).

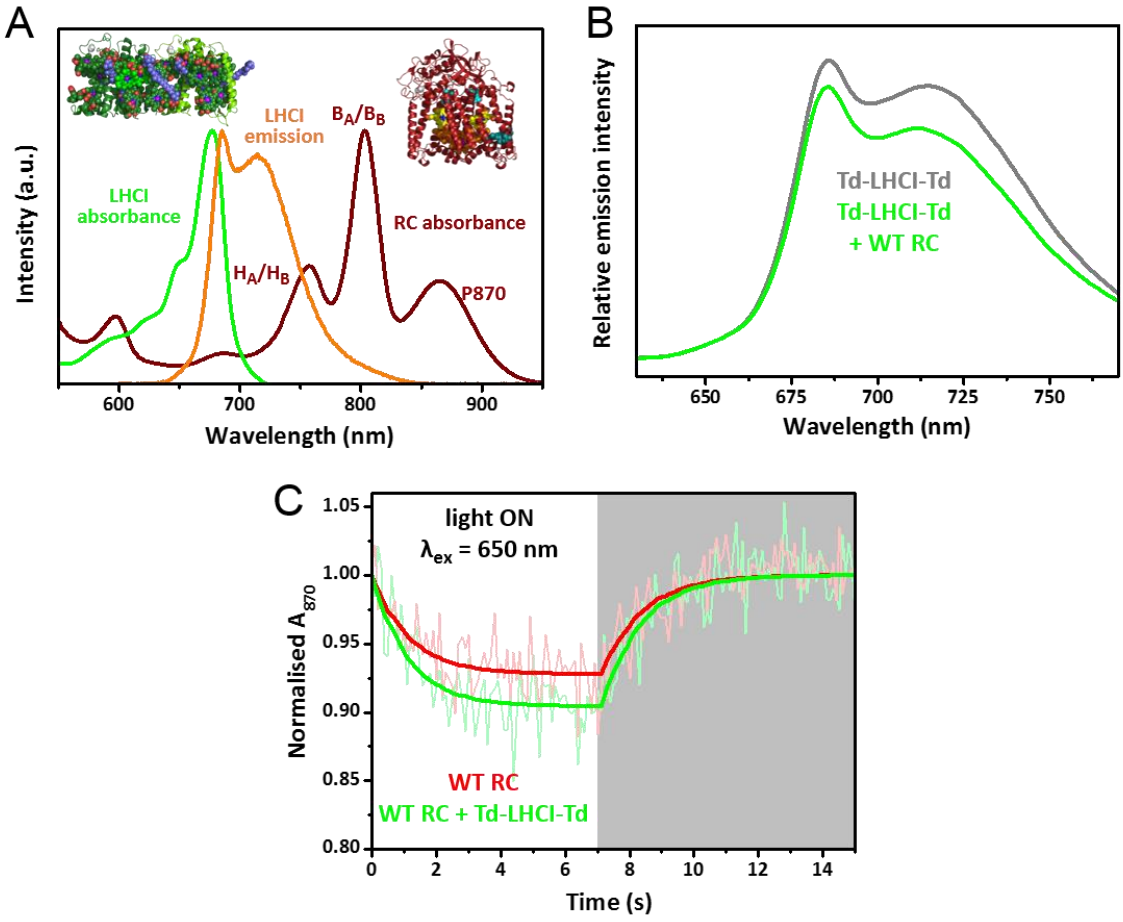


Figure 5.3. Mixing unadapted LHCI and RCs in solution. (A) The red-enhanced emission spectrum of heterodimeric plant LHCI has a stronger overlap with the absorbance spectrum of the *Rba. sphaeroides* RC, particularly the coincident absorbance bands of the BPhes (H_A/H_B). (B) LHCI emission in the absence and presence of WT RCs (excitation at 500 nm). (C) Data and fits for photobleaching and dark recovery of P870 absorbance for the WT RC in the absence and presence of LHCI. Kinetic constants from the fits are summarised in **Table S5.1**. Excitation was for 7s through a ~50 nm wide bandpass filter centred at 650 nm. In all panels the variant of recombinant LHCI heterodimer used had SpyTag attached to the N-terminus of both Lhca1 and Lhca4.

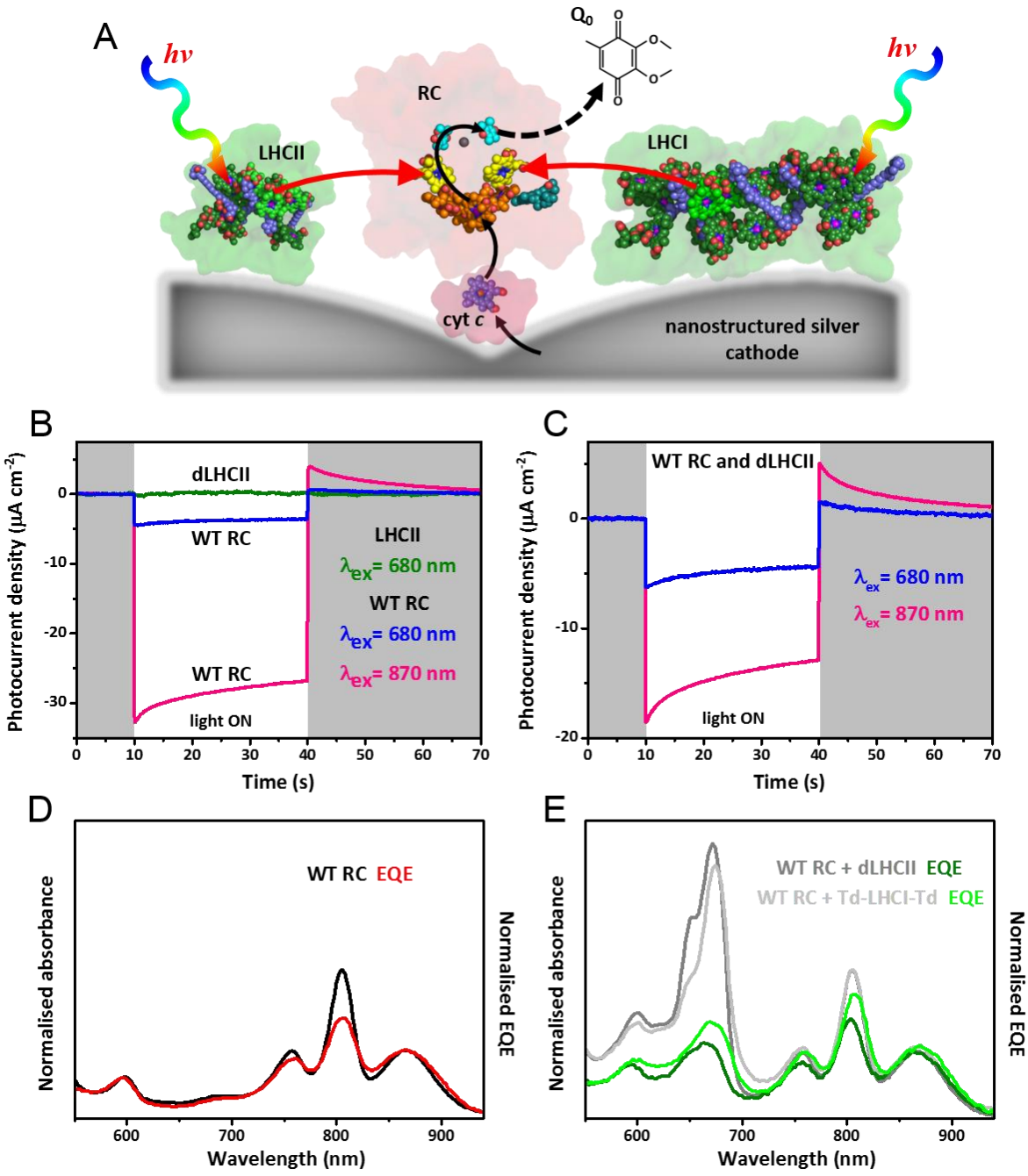


Figure 5.4. Photocurrents from electrodes coated with unadapted proteins. (A) Schematic of photocurrent generation on a nanostructured silver electrode; black arrows show the route of electron transfer, cyan arrows show energy flow. (B) Photocurrents from photocathodes prepared by incubation with 100 µM WT RCs (magenta/blue) or 100 µM dLHCII (green). (C) Photocurrents from a photocathode formed from a 1:1 mix of 100 µM WT RCs and 100 µM dLHCII. (D) Solution absorbance and EQE spectra for WT RCs. (D) Solution absorbance and EQE spectra for 1:1 mixtures of WT RCs and dLHCII or Td-LHCI-Td. The absorbance spectra in (D,E) were normalised at 804 nm and are shown on the same scale, whilst each EQE spectrum was normalised to the corresponding absorbance spectrum at the

maximum of the P870 band. The variant of LHCI used to record the EQE spectrum in (E) is described in Section 5.2.3.

5.2.3. Design and production of components for chimeric photosystems

In an attempt to activate Chl to BChl energy transfer in dilute solution, RCs and LHCs were adapted using the SpyTag/SpyCatcher protein fusion system[147] as a programmable interface (see Section 1.8.4.1 of Chapter 1). When mixed in solution, highly-specific binding of the short SpyTag peptide to the SpyCatcher protein domain initiates autocatalysis of an isopeptide bond between the two involving aspartate and lysine residues (**Fig. 5.1D**), producing a single, covalently-locked, water-soluble protein domain. Using synthetic genes, the SpyCatcher protein was fused to the RC and the short SpyTag peptide to recombinant LHCII or LHCI (see Section 2.4 for the methods used). This created adapted components that could be expressed independently using *Rba. sphaeroides* or *E. coli*, respectively (see Sections 2.5 and 2.6).

To adapt the RC an optimized version of SpyCatcher[156], 106 amino acids in length (SpyCatcherΔ), was fused to the N-terminus of PufL which is an alanine exposed at the protein surface on the cytoplasmic side of the membrane (white spheres in **Fig. 5.1A**). SpyCatcherΔ lacked nine C-terminal amino acids that are not resolved in the X-ray crystal structure of SpyCatcher/SpyTag[156] and was used in order to reduce the length of the linking peptide between the main bodies of the RC and SpyCatcher proteins. The fusion was either direct or via a four (SESG) or eight (SGESGESG) amino acid linker and was preceded by a His-tag for purification (**Fig. 5.5A**). The resulting RCs were named RC*C*, RC*4C* and RC*8C*, respectively, where *C* denotes SpyCatcherΔ and the italicised number the linker (**Fig. 5.5A**). The nomenclature used for these RCs and all other component proteins and chimeras are summarised in **Table 5.1**. Adapted RC proteins were expressed in *Rba. sphaeroides* (Section 2.5), purified and stored in Tris/DDM buffer.

As in Chapter 4, LHCII constructs were derived from a pET-28a vector containing a gene encoding the Lhcb1.3 protein from *A. thaliana* (UniProtKB entry P04778), that was a kind gift from Prof. Roberta Croce, Vrije Universiteit Amsterdam. Modified Lhcb1 apoproteins were expressed in *E. coli* and mature pigment-protein monomers refolded *in vitro* with purified pigments[72,145,179,227]. Four LHCII proteins were designed (**Fig. 5.5B**,**D**). The first had a His-tag at its C-terminus and lacked twelve dispensable N-terminal amino acids that are not resolved in available X-ray crystal structures[64–66] This sequence, which includes four basic amino acids,

is involved in stacking of thylakoid grana but can be removed without affecting core LHCII light-harvesting function[65]. Their removal minimised the sequence linking the main body of LHCII (starting at serine 14 – see **Fig. 5.1B**) to additional components added at the N-terminus. This "dLHCII" protein was used as the control LHCII complex for all measurements, including those described in **Figs. 5.2-5.4**). The remaining three LHCII constructs had either the full 13 amino acid SpyTag peptide or a truncated 10 amino acid SpyTagΔ variant added to the N-terminus of the truncated Lhcb1 (named T-dLHCII and Td-dLHCII, respectively) or SpyTag added to the C-terminus of the full Lhcb1 (named LHCII-T) (**Fig. 5.5B**). The modified SpyTagΔ lacked three dispensable amino acids at its C-terminus[156], further reducing the linker to the N-terminus of LHCII. In all cases a His-tag placed adjacent to the SpyTag sequence ensured the latter was retained in the final, purified pigment-protein (**Fig. 5.5B**). Final holoprotein preparations were stored in Tris/DDM.

The four refolded LHCII complexes had absorbance spectra that were similar to one another (**Fig. 5.6A**) and to spectra previously published by others[72,145,179,227]. Their emission spectra were highly similar (**Fig. 5.6B**), and the line-shapes of these spectra were invariant with excitation wavelength (shown in **Fig. S5.6C** for dLHCII), a feature diagnostic of a structurally-intact LHCII. Pigment compositions were similar to those typically reported for recombinant LHCII (**Fig. 5.6D**)[145].

Adapted LHCI heterodimeric complexes were assembled from modified versions of the *A. thaliana* proteins Lhca1 (UniProtKB entry Q01667) and Lhca4 (UniProtKB entry P27521). The mature Lhca1 was modified at its N-terminus either with a Myc protein purification affinity tag (named L1) or with a Myc-tag followed by the shortened ten amino acid SpyTagΔ (named Td-L1), and the mature Lhca4 was modified at its N-terminus with a His-tag followed by SpyTagΔ (named Td-L4) (**Fig. 5.5C,D**). The expression plasmids was pET-28a and synthetic LHCI genes were sourced from Eurofins.

LHCI heterodimeric pigment-proteins were also refolded from apoproteins expressed in *E. coli* [73,76,180,240]*.* This involved mixing the SpyTagΔ-adapted Lhca4 protein (Td-L4) with a 20 % excess (by mass) of either unadapted Lhca1 protein (L1) or SpyTagΔ-adapted Lhca1 protein (Td-L1), to produce LHCI either singly or doubly modified with SpyTagΔ (named LHCI-Td and Td-LHCI-Td, respectively). This enabled the creation of chimeras between the LHCI heterodimer and either one or two RCs. Nickel affinity chromatography was used to separate the His-tagged LHCI dimer from residual Lhca1 monomer (which was not His-tagged). The final LHCI-Td and Td-LHCI-Td proteins were stored in Tris/DDM and were used in the

measurements described in **Figs. 5.3** and **5.4** to look at interactions of LHCI with unadapted RCs.

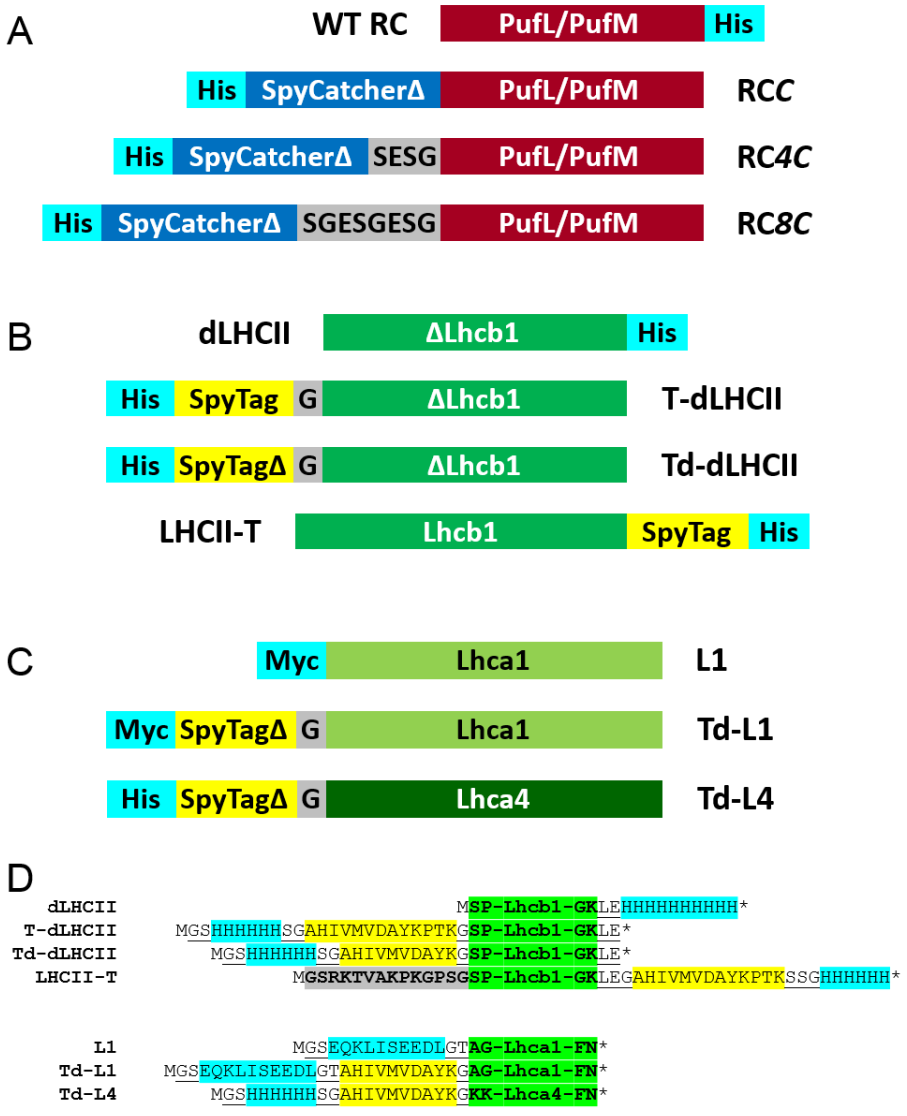
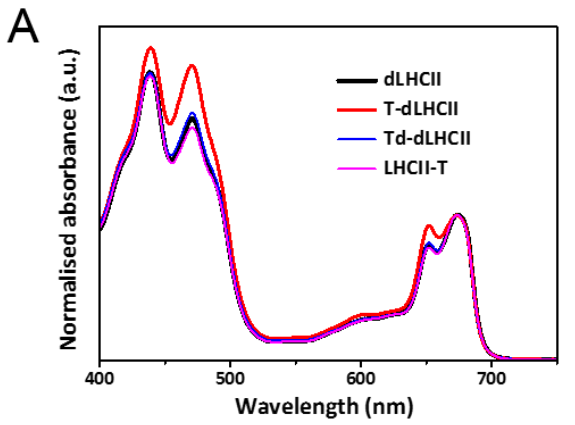
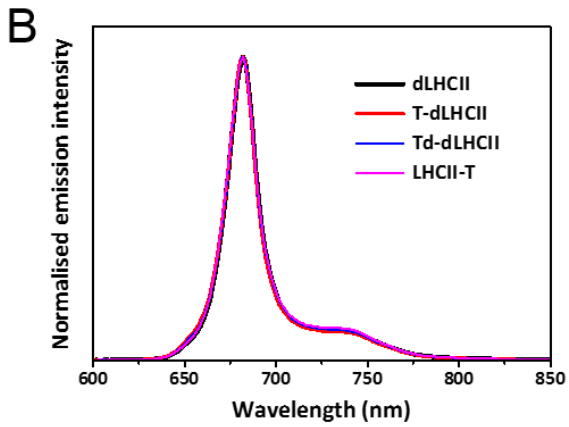


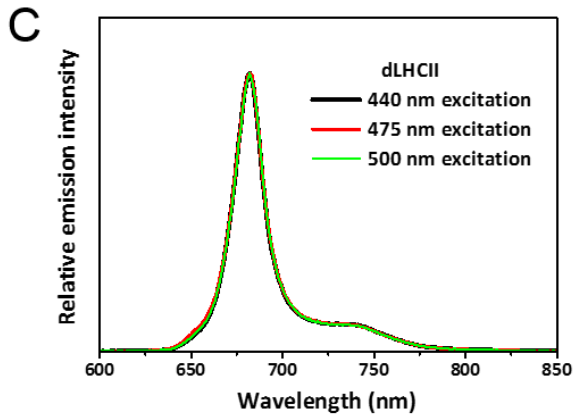
Figure 5.5. Design of components for RC#LHC chimeras. (A) Construct designs for adaptation of the RC. For purification the WT RC was modified with a His-tag on PufM. (B) Construct designs for adaptation of LHCII. The control LHCII was truncated at its N-terminus (dLHCII – see text) and was His-tagged at its C-terminus. (C) Construct designs for adaptation of LHCI which is a Lhca1/Lhca4 heterodimer. (D) LHC protein sequences are marked in bold/green highlight, with the dispensable sequence at the N-terminus of Lhcb1 highlighted in grey. Added His and Myc (EQKLISEEDL) tags are highlighted in cyan, the full-length and truncated version of SpyTag are highlighted in yellow, and linkers are underlined. Only the first and last two residues of each LHC protein are shown.

Table 5.1. Nomenclature and descriptions of pigment-protein complexes.

Name	Description
SpyCatcherΔ	SpyCatcher lacking nine C-terminal amino acids
SpyTagΔ	SpyTag lacking three C-terminal amino acids
RC*C*	SpyCatcherΔ attached to N-terminus of RC PufL
RC*4C*	SpyCatcherΔ attached to N-terminus of RC PufL by a 4 amino acid linker
RC*8C*	SpyCatcherΔ attached to N-terminus of RC PufL by a 8 amino acid linker
dLHCII	Lhcb1 lacking 12 N-terminal amino acids, His tag at C-terminus
T-dLHCII	His-tag and full SpyTag added to N-terminus of dLHCII
Td-dLHCII	His-tag and truncated SpyTagΔ added to N-terminus of dLHCII
LHCII-T	Full SpyTag and His-tag added to C-terminus of Lhcb1
LHCII#RC	Chimera of LHCII-T and RC*C*
RC*4*#dLHCII	Chimera of RC*4C* and Td-dLHCII
Td-L4	His-tag and SpyTagΔ added to N-terminus of Lhca4
L1	Myc-tag added to N-terminus of Lhca1
Td-L1	Myc-tag and SpyTagΔ added to N-terminus of Lhca1
LHCI-Td	LHCI heterodimer from Td-L4 and L1
Td-LHCI-Td	LHCI heterodimer from Td-L4 and Td-L1
LHCI#RC	Chimera of LHCI-Td and one RC*C*
RC#LHCI#RC	Chimera of Td-LHCI-Td and two RC*C*







D

Sample	Chl *a*/Chl *b*	Chl/[LHCII]
dLHCII	1.43	12.51
T-dLHCII	1.09	13.48
Td-dLHCII	1.37	12.85
LHCII-T	1.50	11.96

Figure 5.6. LHC adaptation and LHCII refolding. (A) Absorbance spectra of the four refolded LHCII variants, normalised at the maximum of the low energy chlorophyll *a* band. (B) Normalised emission spectra of these complexes obtained with 440 nm excitation. (C) Normalised emission spectra for dLHCII illustrating that emission profile was independent of excitation wavelength. (D) Deduced cofactor compositions for each refolded LHCII.

5.2.4. Self-assembly of two-component RC-LHCII chimeras

When incubated together, all possible pairwise combinations of the three SpyCatcherΔ-adapted RCs and three SpyTag-adapted LHCIIs efficiently formed a stable higher molecular weight product, dubbed a chimera. Covalent-locking of the structure through the SpyCatcher/SpyTag module enabled purification of all nine LHCII-RC chimeras with the designed molar ratio by gel filtration chromatography (see two examples in **Fig. S5.2A**).

The mass change associated with chimera formation could be visualised on linear sucrose density gradients prepared by freezing and thawing 21 % (w/v) sucrose in 20 mM Tris/0.04% DDM (pH 8.0). Unadapted RCs and smaller unadapted LHCIIs could be distinguished as discrete red and green bands (**Fig. 5.7A**, gradients 1,2), as could components in mixtures incapable of forming chimeras (**Fig. 5.7A**, gradients 3,4). However, mixing any SpyCatcherΔ-adapted RC with any SpyTag-adapted LHCII produced a product that migrated further than either monomeric protein. The two examples shown (**Fig. 5.7A**, gradients 5,6) are LHCII#RC chimeras from a RC*C*/LHCII-T mix and RC*4*#dLHCII chimeras from a RC*4C*/Td-dLHCII mix, where "#" denotes the spontaneously-formed SpyCatcher/SpyTag interface domain (**Table 5.1**).

Chimera formation could also be detected on native blue gels (**Fig. 5.8A**, bands at LHCII-#-RC arrow labelled as PRODUCT), and SDS-PAGE combined with western blotting using anti-His antibodies confirmed that chimera self-assembly was due to the formation of a covalent fusion protein between the SpyTag-adapted or SpyTagΔ-adapted Lhcb1 polypeptide of LHCII and the SpyCatcherΔ-adapted PufL polypeptide of the RC (**Fig. 5.8B**, Lhcb1-#-PufL band labelled as PRODUCT).

A change in protein morphology on chimera formation was also evident from TEM. Images of a mix of unadapted WT RCs and dLHCII showed a large number of monodispersed, regularly-sized objects of <10 nm diameter (**Fig. 5.7B**, left), whereas images of the purified LHCII#RC chimera revealed larger, evenly-distributed elliptical structures of >10 nm diameter along the longest axis (**Fig. 5.7B**, right). A molecular model of this chimera, based on available X-ray crystal structures for the RC, LHCII and SpyCatcher/Tag, is shown in **Fig. 5.7C**.

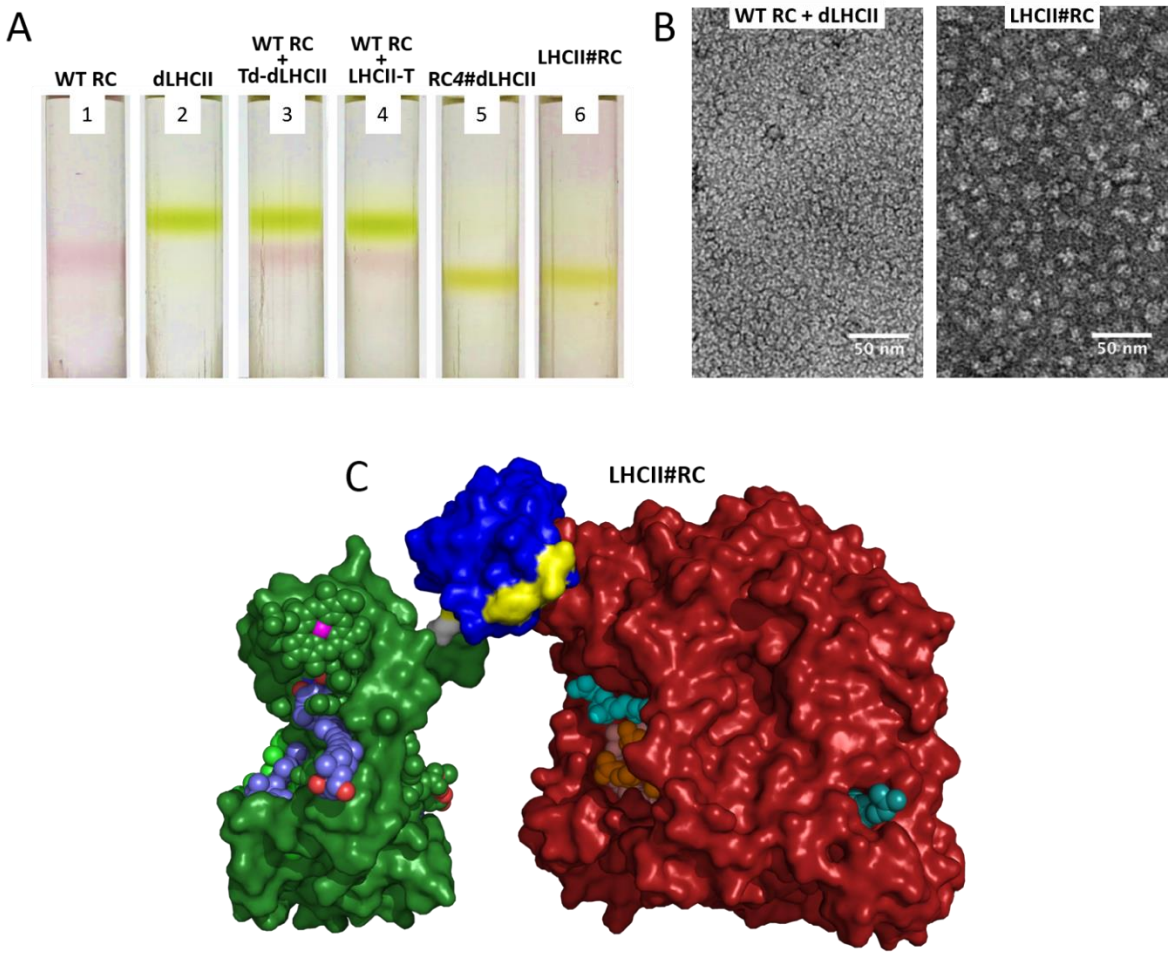


Figure 5.7. Assembly of RC#LHCII chimeras. (A) Sucrose density gradient fractionation of RCs (red bands) and LHCII complexes (green bands) on linear sucrose gradients were prepared by freezing and thawing 10 mL of 21 % (w/v) sucrose in 20 mM Tris/0.04% DDM (pH 8.0). Each gradient was loaded with 400 μL of sample with each photoprotein at a concentration of 2.5 μM and then capped with 1 mL of 20 mM Tris/0.04% DDM (pH 8.0). Gradients were ultracentrifuged at 38,000 rpm for 18 hours at 4 °C. RC-LHCII chimeras migrate to a lower position in gradients than either RC or LHCII monomers, with no dissociation into components. (B) TEM images of (left) an equimolar mixture of the WT RC and dLHCII and (right) the purified LHCII#RC chimera. (C) Molecular model of the LHCII#RC chimera produced using Modeller[189]. The RC (maroon) N-terminally adapted with SpyCatcherΔ (blue) is covalently linked to LHCII (green) C-terminally adapted with SpyTag (yellow). Cofactor colours are as described in **Fig. 5.1**. Protein structures used were Protein Data Bank entries 3ZUW for the *Rba. sphaeroides* RC[241], 2BHW for the LHCII from pea[65] and 4MLI for SpyCatcher/Tag[156].

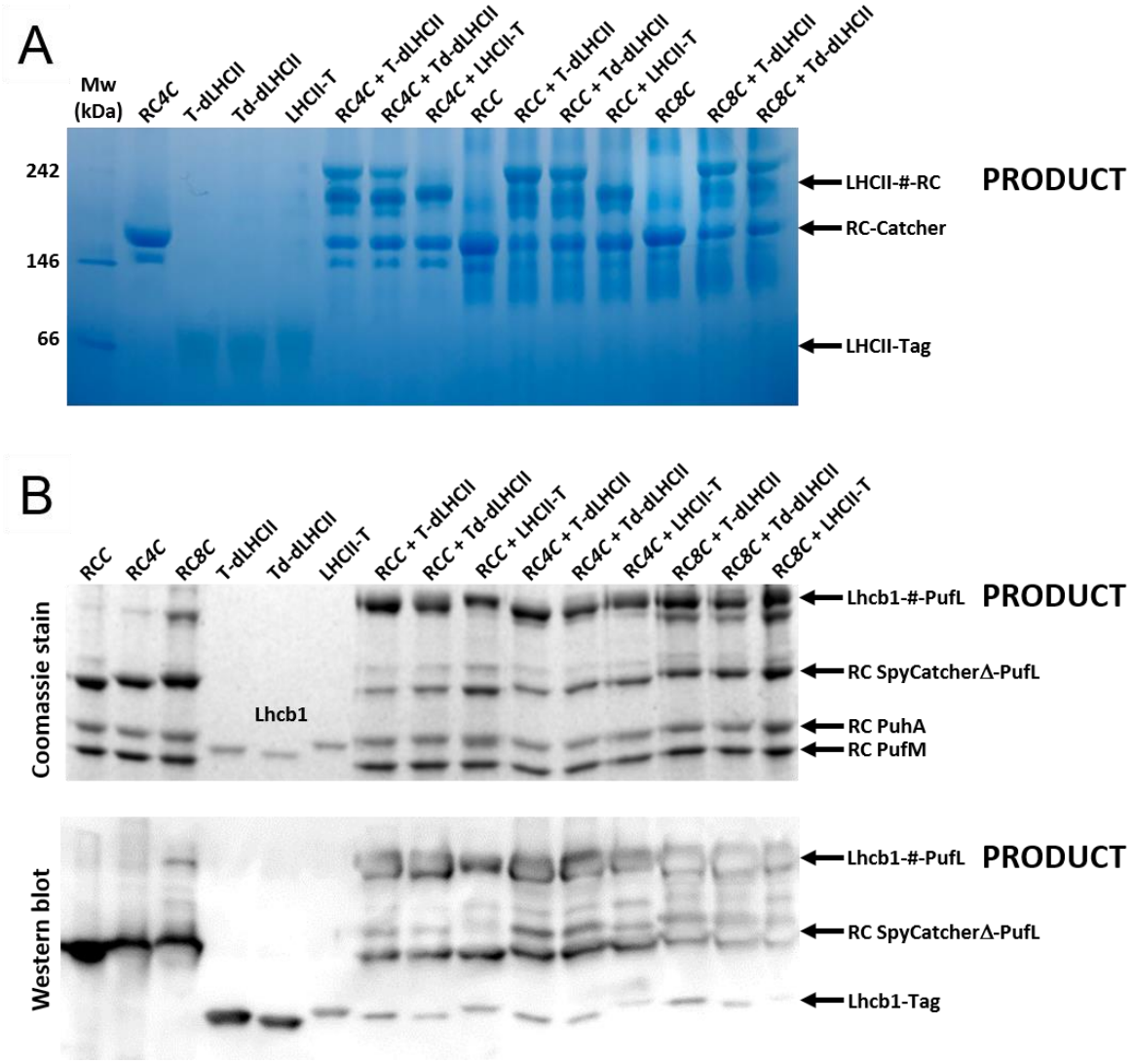


Figure 5.8. Formation of covalent RC#LHCII chimeras. Gels and blots show the outcome of mixing a two-fold excess of adapted RCs with adapted LHCIIs. (A) Blue native PAGE of individual proteins and mixtures. The combination LHCII-T and RC8*C* was not run in this experiment. Mixing of SpyCatcherΔ-adapted RCs (bands labelled RC-Catcher) with SpyTag or SpyTagΔ-adapted LHCII complexes (bands labelled LHCII-Tag) produces higher molecular weight products denoted as LHCII-#-RC (labelled PRODUCT) where # is the SpyCatcher/Tag domain. (B) SDS PAGE and anti-His western blotting of individual proteins and mixtures. The RC is made up from three polypeptides, PufL (L-polypeptide), PufM (M-polypeptide), PuhA (H-polypeptide). Mixing of SpyCatcherΔ-adapted RCs with SpyTag- or SpyTagΔ-adapted LHCII complexes produces a covalently-linked Lhcb1-#-PufL polypeptide (labelled PRODUCT) where # is the SpyCatcher/Tag sequences connected by an isopeptide bond. Western blotting detected the His-tag on the Lhcb1-#-PufL fusion peptide and the His-tags on unreacted RC SpyCatcherΔ-PufL and residual Lhcb1-Tag variants.

5.2.5. Self-assembly of two- and three-component RC-LHCI chimeras

RC-LHCI and $(RC)_2$-LHCI chimeras were assembled by incubation of LHCI-Td or Td-LHCI-Td with a three-fold excess of RC*C*. Chimera formation produced higher molecular weight products that could be separated from unreacted RCs on blue native gels (see examples in **Fig. 5.9A**). As designed, assembly of RC*C* with doubly-adapted Td-LHCI-Td complexes produced higher molecular weight products than with singly-adapted LHCI-Td complexes (**Fig. 5.9A**, right).

Equivalent results were obtained with LHCI adapted with the full SpyTag and also with RC*4C* (**Fig. 5.10A**, band labelled PRODUCT). Analysis by SDS-PAGE and western blotting confirmed that chimera self-assembly was due to spontaneous formation of fusion proteins between the SpyCatcherΔ-adapted PufL of the RC and the SpyTagΔ-adapted Lhca4 and Lhca1 (**Fig. 5.10B**, bands labelled PRODUCT).

Chimeras formed from RC*C* were selected for further analysis as these had no linker between the SpyCatcher/Tag domain and the RC. The chimera formed between LHCI-Td and one RC*C* was named LHCI#RC and that formed between Td-LHCI-Td and two RC*C* was named RC#LHCI#RC. These were purified with the designed molar ratio by gel filtration chromatography (**Fig. S5.2B**).

Sucrose density gradient ultracentrifugation (**Fig. 5.9B**) showed that LHCI#RC chimera (gradient 5) was clearly larger than LHCI alone (gradients 2-4) or unadapted RCs (gradients 1,3,4), and the RC#LHCI#RC chimera (gradient 6) was larger again. Molecular models of these chimeras are shown in **Fig. 5.9CD**.

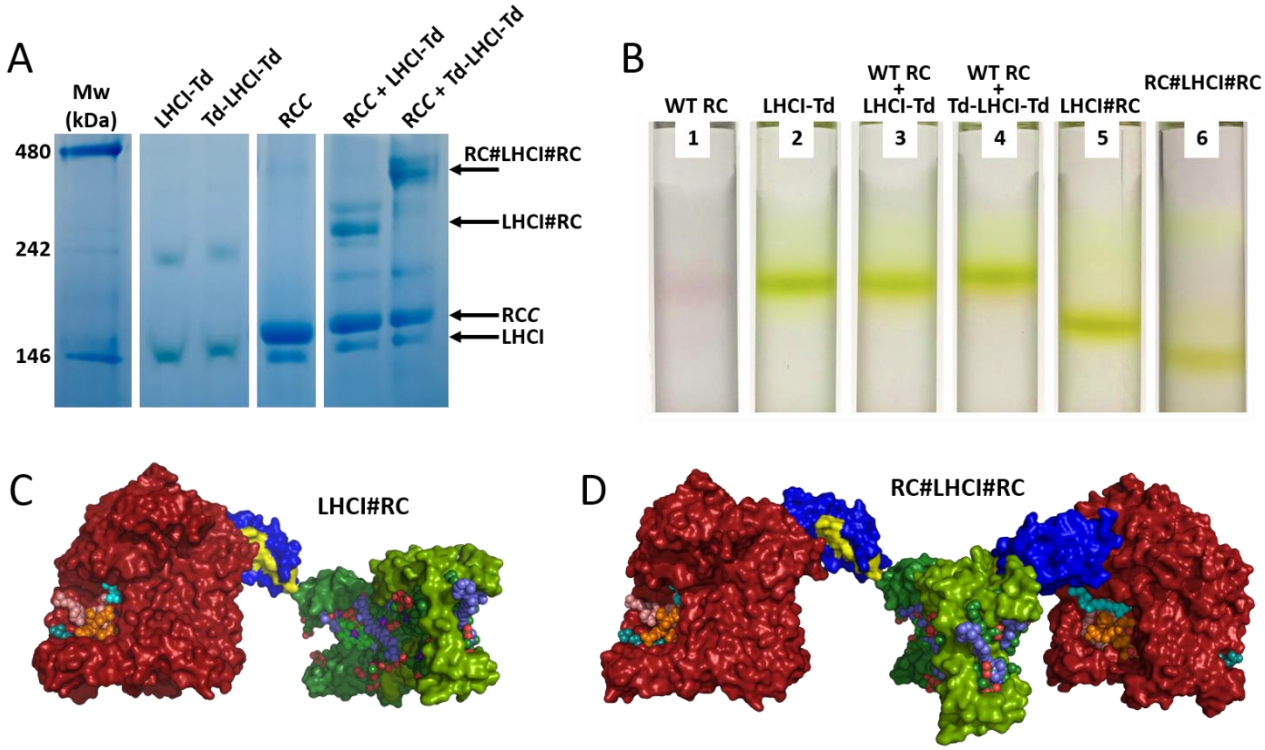


Figure 5.9. Engineering of RC#LHCI chimeras. (A) Blue native PAGE showing the formation of high molecular weight products by mixing LHCI-Td or Td-LHCI-Td with RC*C* (see **Fig. 5.10A** for the full gel with other combinations). The multiple bands seen for the high molecular weight products are likely to be due to conformational heterogeneity. (B) Sucrose density gradient fractionation of RCs (red bands) and LHCI (green bands). LHCI#RC chimeras and larger RC#LHCI#RC chimeras migrate to lower positions than either RCs or LHCI. (C,D) Molecular models of the LHCI#RC (C) and RC#LHCI#RC (D) chimeras produced using Modeller[189]. Cofactor colours are as described in **Fig. 5.1**. Protein structures used were Protein Data Bank entries 3ZUW for the *Rba. sphaeroides* RC[241], 4KX8 for the LHCI from pea[29] and 4MLI for SpyCatcher/Tag[156].

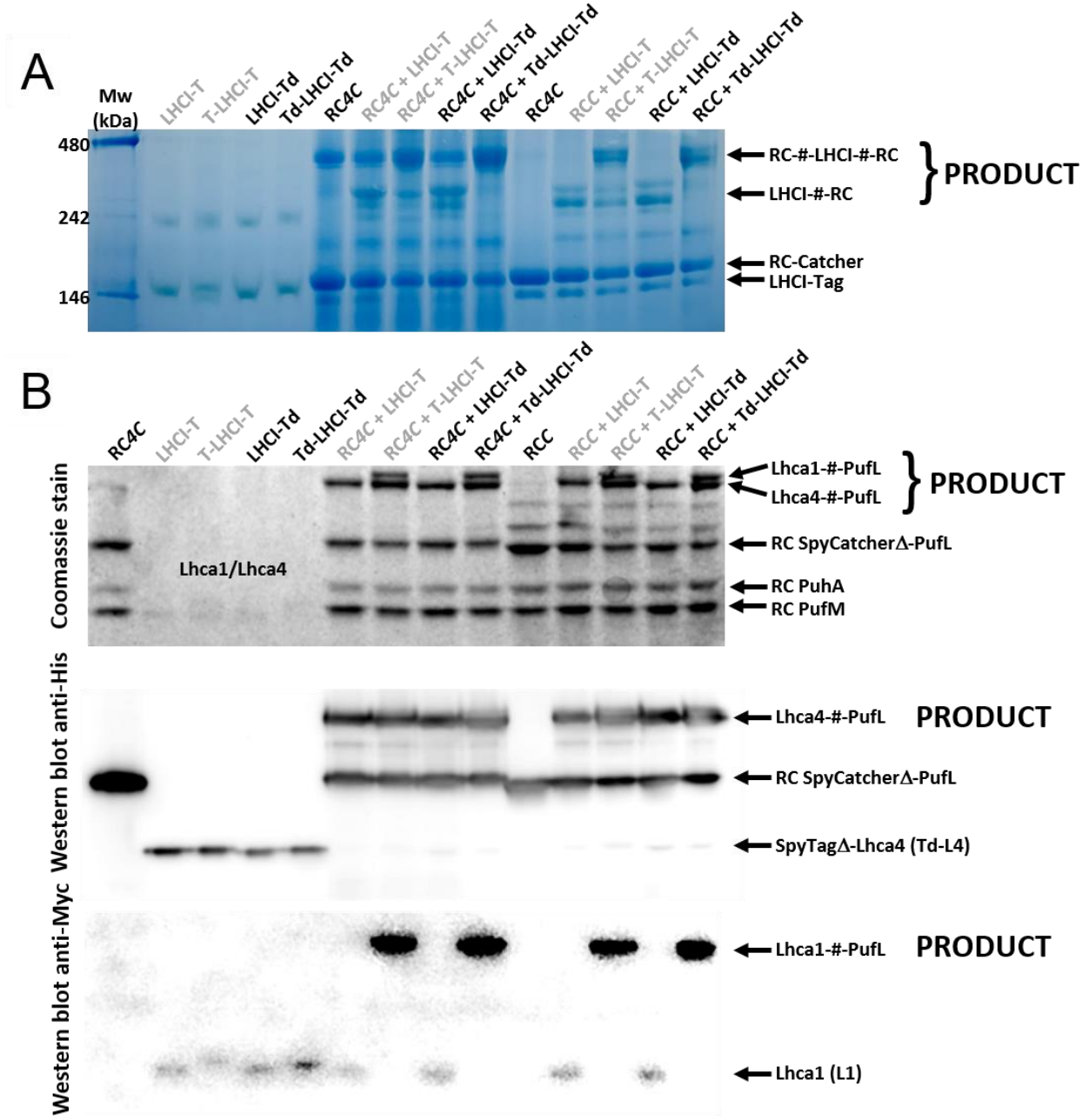


Figure 5.10. Formation of covalent RC#LHCI chimeras. Gels and blots show the effect of mixing a three-fold excess of adapted RCs with single or doubly adapted LHCI complexes. In addition to combinations described in the text, data are shown for an equivalent set of constructs in which LHCI was either singly (LHCI-T) or doubly (T-LHCI-T) modified with the full length SpyTag peptide (lane labels shown in grey), which produced identical results. (A) Blue native PAGE of individual proteins and mixtures. Mixing of SpyCatcherΔ-modified RCs (labelled RC-Catcher) with singly or doubly SpyTag- or SpyTagΔ-modified LHCI complexes (labelled LHCI-Tag) produces higher molecular weight products denoted as LHCI-#-RC or RC-#-LHCI-#-RC (labelled PRODUCT) where # is the SpyCatcher/Tag domain. The multiple

bands seen for the high molecular weight products may have arisen from conformational heterogeneity. (B) SDS PAGE and western blotting of individual proteins and mixtures revealing formation of covalently-bound Lhca4-#-PufL and/or Lhca1-#-PufL fusion proteins. LHCI is a heterodimer of Lhca1 and Lhca4. Mixing of SpyCatcherΔ-adapted RCs with singly SpyTag- or SpyTagΔ-adapted LHCI complexes (on Lhca4) produces a covalently linked Lhca4-#-PufL polypeptide where # is the SpyCatcher/Tag sequences connected by an isopeptide bond. Mixing of SpyCatcherΔ-adapted RCs with doubly SpyTag- or SpyTagΔ-modified LHCI complexes produces covalently-linked Lhca1-#-PufL and Lhca4-#-PufL polypeptides (labelled PRODUCT). Western blotting with anti-His antibodies detected the His-tag on the Lhca4-#-PufL fusion peptide and the His-tags on unreacted RC SpyCatcherΔ-PufL and residual Lhca4 Td-L4 polypeptide. Western blotting with anti-Myc antibodies detected the Myc-tag on the Lhca1-#-PufL fusion peptide and the Lhca1 L1 polypeptide (the weak reaction to the latter relative to the large fusion peptide was probably due to inefficient transfer to the blot membrane).

5.2.6. Chl to BChl energy transfer in chimeras in dilute solution

Fluorescence spectroscopy was used to look for evidence of energy transfer in the RC-LHC chimeras in dilute solution. The nine chimeras formed between RCs and LHCIIs exhibited between 14 and 27% less LHCII emission than a control sample formed from an equivalent mix of the SpyTag-adapted LHCII and WT RCs (**Fig. 5.11A** – and see spectra in **Fig. S5.3A**). This was diagnostic of energy transfer, probably through FRET at the distances implied by the chimera models (**Fig. 5.7C**), that was activated by physically-linking the RC to the LHCII. For the chimeras based on LHCIIs with SpyTag at the N-terminus (**Fig. 5.11A**, left and centre) the extent of emission quenching did not seem to depend on the type of SpyCatcher-adapted RC, but for those based on the LHCII with SpyTag at the C-terminus the quenching seemed to be greater for RC*s* with a shorter linker to the SpyCatcher module (**Fig. 5.11A**, right).

These trends, observed with 651 nm excitation, were also seen in data on the same complexes gathered with 475 nm excitation (**Fig. S5.3A**, left and **S5.3B**). This lack of an excitation wavelength dependence showed that the LHCIIs in the chimeras were structurally intact and that the reduction in their LHCII emission was not due to parasitic RC absorbance, which would be expected to be wavelength dependent.

To confirm energy transfer from LHCII to the RC, measurements of RC P870 photo-oxidation in response to 650 nm excitation were carried out on LHCII#RC and RC*4*#dLHCII

chimeras as examples where the LHCII was connected via either its N- or C-terminus. Bleaching of 870 nm absorbance was much stronger in LHCII#RC chimeras than in controls comprising the RC*C* alone, or a mixture of RC*C* with unadapted dLHCII complexes (**Fig. 5.11B**), and the same was found for the RC*4*#dLHCII chimera (**Fig. 5.11C**). Hence decreased emission by the LHCII energy donor was accompanied by enhanced photo-oxidation of the RC energy acceptor, confirming energy transfer between the two in solution.

With LHCI, greater quenching of emission was seen on forming LHCI#RC or RC#LHCI#RC chimeras with either RC*C* or RC*4C* than on mixing the same adapted LHCI with unadapted WT RCs (**Fig. 5.12A**). These findings were again independent of excitation wavelength (see spectra and histograms in **Fig. S5.4**) showing they were not due to absorbance of excitation light by the tethered RC. This emission quenching was accompanied by significant enhancement of P870 photo-oxidation in LHCI chimeras with one or two RC*C*, compared to RC*C* alone (**Fig. 5.12B**), confirming energy transfer from LHCI to the RC in solution on formation of a chimera between the two.

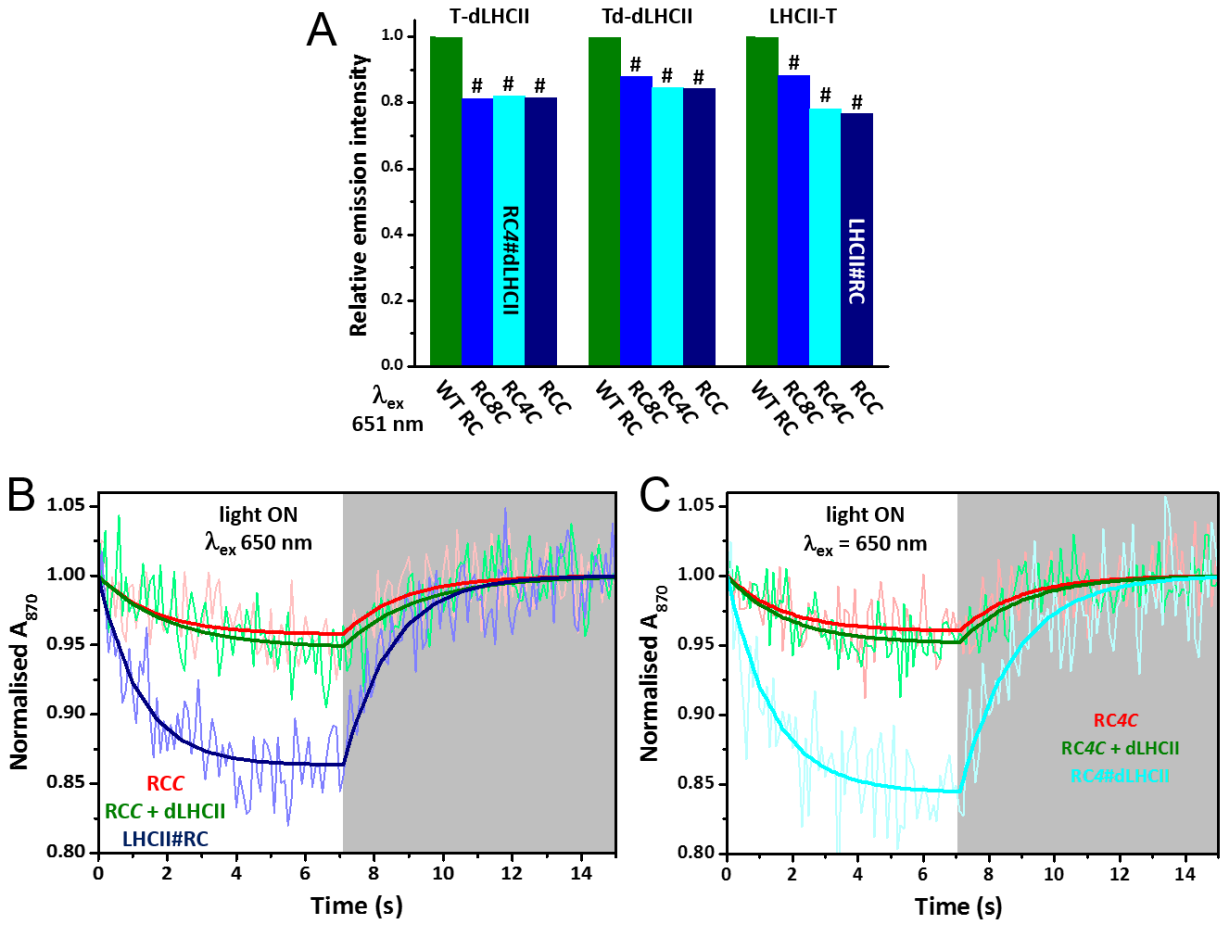


Figure 5.11. Energy transfer in RC#LHCII chimeras in solution. (A) LHCII emission at 681 nm from mixtures with a 2:1 RC:LHCII composition. Emission from LHCII in chimeras (shades of blue) is expressed relative to that from the same LHCII mixed with WT RCs (green). Symbol # denotes mixtures where LHCII has self-assembled into chimeras. Mixtures containing the two chimeras analysed in depth are labelled. (B) Data and fits for photobleaching and dark recovery of P870 absorbance in RC*C*, a 1:1 RC*C*:dLHCII mixture and the LHCII#RC chimera. (C) Data and fits for photobleaching and dark recovery of P870 absorbance in RC*4C*, a 1:1 RC*4C*:dLHCII mixture and the RC*4*#dLHCII chimera.

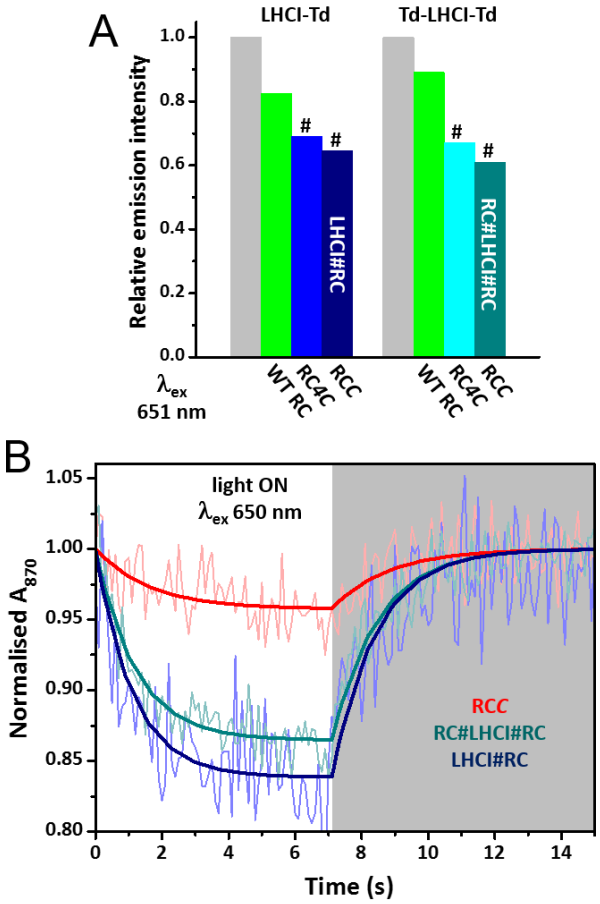


Figure 5.12. Energy transfer in RC#LHCI chimeras in solution. (A) LHCI emission at 714 nm from mixtures with a 3:1 RC:LHCI composition. Emission from RC-LHCI chimeras (left) or $(RC)_2$-LHCI chimeras (right), and that from a mixture of the same LHCI complex and WT RCs (green) is expressed relative to that from the LHCI complex alone (grey). Symbol # denotes mixtures where LHCI has self-assembled into a chimera. Mixtures containing the two chimeras analysed in depth are labelled. (B) Data and fits for photobleaching and dark recovery of P870 absorbance in RC*C* and the two RC-LHCI chimeras.

5.2.7. Functionality of chimeras on electrodes

One motivation to create an all-protein tethering system as opposed to using a QD hub was to create a smaller and less complex polychromatic hybrid with better access to the RC for electrodes and mediators. To test their functionality, purified chimeras were adhered to nanostructured silver electrodes. Promisingly, all those tested were able to generate photocurrents, showing that dynamic interactions between the RC, cyt *c* and ubiquinone at the electrode surface, required for the generation of a photocurrent, were not obstructed by attaching the RC to LHCII or LHCI. All EQE action spectra recorded for LHCII-RC chimeras (**Fig. 5.13A**) or LHCI-RC chimeras (**Fig. 5.13A**) exhibited the low energy Q_y Chl band indicating photocurrent generation powered by LHC absorbance. EQE spectra recorded using a xenon source also showed the high energy LHCII or LHCI Chl Soret band (**Fig. S5.1DE**).

The probable mechanism of solar energy conversion operating in chimeras, based on what is known about the photophysics of these proteins, is summarised in **Fig. 5.14**. Energy captured by the pigment systems of LHCII or LHCI is passed to the RC, exciting the primary electron donor BChls (P870*) and initiating charge separation to form $P870^{+}Q_B^{-}$. Energy harvested by the Chl *b* (or carotenoid) pigments of either LHC is passed to the lower energy Chl *a*. Inter-protein energy transfer is likely to involve a sub-set of red-shifted Chl *a* in either LHC and entry into the RC is likely to occur principally via the BPhe cofactors ($H_{A/B}$) as their absorbance has the greatest spectral overlap with LHC emission (**Figs. 5.2A** and **5.3A**). On an electrode with suitable mediators charge separation in the RC initiates the generation of a photocurrent.

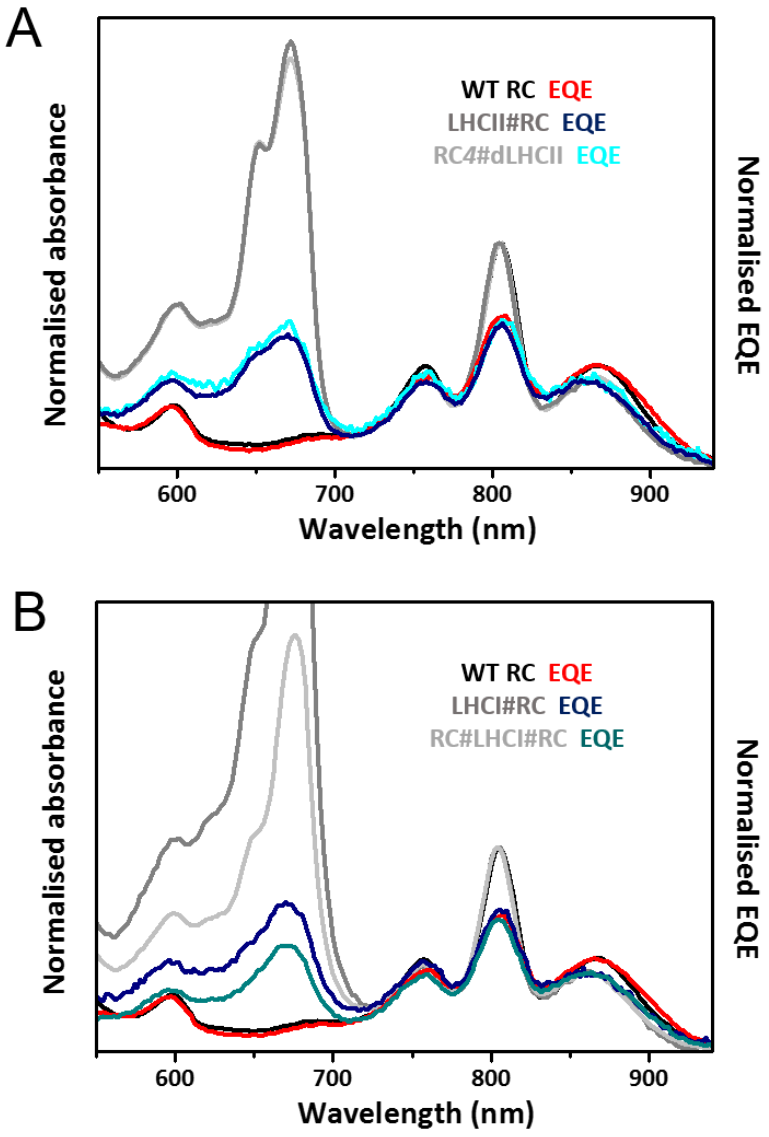


Figure 5.13. Energy transfer in chimeras on an electrode. (A) Solution absorbance and EQE spectra for WT RCs compared with those for LHCII#RC and RC*4*#dLHCII chimeras. (B) Solution absorbance and EQE spectra for WT RCs compared with those for LHCI#RC or RC#LHCI#RC chimeras. For both panels the absorbance spectra were normalised at 804 nm, whilst each EQE spectrum was normalised to the corresponding absorbance spectrum at the maximum of the P870 band.

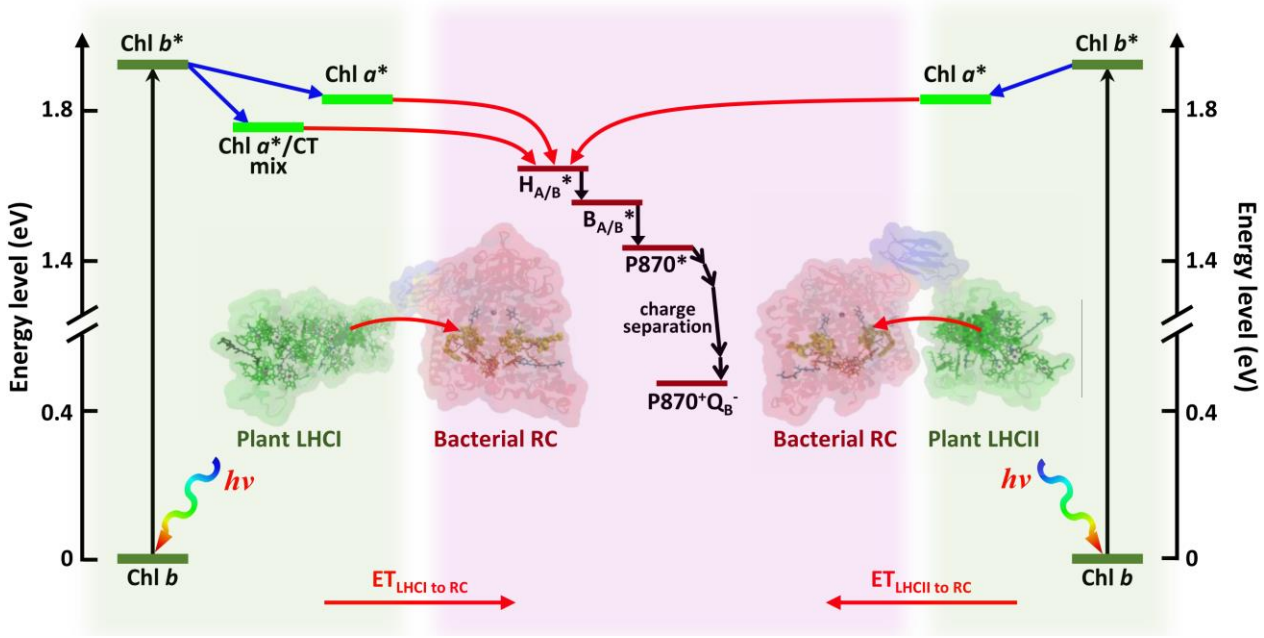


Figure 5.14. Solar energy conversion in chimeras. Energy flow within LHCII or LHCI is from higher energy Chl *b* to lower energy Chl *a*. LHCI also exhibits a red-shifted emissive state with mixed excitonic/charge transfer (CT) character. Excited state energy entering the RC via the BPhes ($H_{A/B}$) migrates to the P870 BChls via the monomeric BChls ($B_{A/B}$), initiating charge separation to form $P870^{+}Q_{B}^{-}$. Energy harvested by the carotenoid pigments of LHCII or LHCI (not shown) is transferred to the RC via their Chls.

5.2.8. Energy transfer efficiency in chimeras in solution

Apparent efficiencies of energy transfer in chimeras in solution were calculated from LHC emission spectra (E_{FL}) using:

$$E_{FL} = 1 - \frac{FL_{chimera}}{FL_{WTRC+LHC}} \quad \textbf{(5.1)}$$

where $FL_{chimera}$ was the intensity of LHC emission in a chimera and $FL_{WTRC+LHC}$ was that in a concentration-matched mixture of the appropriate LHCII or LHCI variant and the WT RC. The same approach was used for estimating the apparent energy transfer efficiency in the mixtures of WT RCs and LHCII or LHCI, expressing $FL_{WTRC+LHC}$ as a function of the emission from the same concentration of the LHC alone (FL_{LHC}). For LHCII, where the line shape of the emission spectrum did not vary as it is a single quantum system[67], maximum emission values were used as a simple measure of emission intensity. For LHCI, which has multiple distinct emission states[76], values of emission intensity (FL_{int}) were produced by integration across the emission spectrum using **Eq. 5.2**, and then applied in **Eq. 5.1**.

$$FL_{int} = \int \frac{\lambda}{hc} FL_{LHCI}(\lambda) d\lambda \quad \textbf{(5.2)}$$

The resulting values for E_{FL} are compiled in **Table 5.2**.

Apparent efficiencies of energy transfer were also calculated from the rate of P870 photobleaching (k_f) from the kinetic analyses summarised in **Table S5.1**. To enable this the intensity of the 650 nm excitation light used in these experiments was kept low such that no more than ~15 % of P870 was oxidised within the lifetime of $P870^+$ (~ 1s), ensuring that photooxidation directly represented the quantity of energy received by either direct absorption by the RC or energy transfer from the tethered LHC. The apparent efficiency of energy transfer (E_{P870}) was estimated from the rate of P870 photobleaching using:

$$E_{P870} = \frac{k_f(chimera) - k_f(RC)}{k_f(RC)} \frac{\int P(\lambda)(1 - 10^{Abs_{RC}(\lambda)}) d\lambda}{\int P(\lambda)(1 - 10^{Abs_{LHC}(\lambda)}) d\lambda} \quad \textbf{(5.3)}$$

where k_f was the rate of P870 oxidation in a chimera (*chimera*) or the equivalent RC-only control (*RC*) (compiled in **Table S5.1**). Integration of incident photon flux (*P*) and the 1-transmission of RCs or LHCs as a function of wavelength provided the number of photons absorbed by either RCs or LHCIIs per unit area per second (**Table S5.1**).

Table 5.2. Apparent energy transfer efficiencies and associated parameters.

System	E_{P870} (%)	E_{FL} (%)	Relative J[a]	Relative quantum yield[b]
dLHCII				1.00
WT RC + dLHCII	0.8 ± 0.1[c]	2.8 ± 1.1	1.00	
Td-dLHCII				1.02
WT RC + Td-dLHCII	1.9 ± 0.3[c]	0.1 ± 1.4	0.99	
RC*4*#dLHCII	21.7 ± 5.1	17.8 ± 0.2		
LHCII-T				1.04
LHCII-T + WT RC	1.2 ± 0.2[c]	-1.6 ± 1.0	1.00	
LHCII#RC	23.1 ± 4.8	23.8 ± 0.8		
LHCI-Td				0.30
LHCI-Td + WT RC	1.3 ± 0.3[c,d]	13.4 ± 2.0[d]	1.80	
LHCI#RC	20.1 ± 4.4	19.7 ± 1.1		
Td-LHCI-Td				0.28
$(WT\ RC)_2$ + Td-LHCI-Td	3.8 ± 0.6[c,d]	11.9 ± 4.5[d]	1.83	
RC#LHCI#RC	29.2 ± 5.9[e]	27.3 ± 3.6[e]		

[a] Spectral overlap of LHC emission with RC molar absorbance, normalised to that of dLHCII and RC (SD < 0.01).

[b] Quantum yield relative to that of dLHCII, obtained by comparing integration of LHC emission (SD in the range 0.01 ~ 0.02).

[c] These low apparent energy transfer efficiencies may have arisen from some reabsorption of LHC fluorescence by RCs. In accord with this the maximum percentage of LHC emission that could be reabsorbed by unconjugated RCs in the P870 bleaching measurements was estimated to be ~1-2 % for LHCII and ~1-3 % for LHCI, based on their absorbance.

[d] The variance in these two pairs of estimates is attributed to E_{FL} largely reflecting a decrease in LHCI quantum yield on adding WT RCs rather than being due to energy transfer (see text).

[e] RCs conjugated to each of Lhca1 and Lhca4 in the LHCI heterodimer.

Values of E_{P870} calculated from experimental data are shown in **Table 5.2**. The efficiency of energy transfer was less than 4 % in mixtures of unadapted RCs with unadapted dLHCII, SpyTag-adapted LHCII or SpyTag-adapted LHCI, consistent with expectations for a dilute (500 nM) solution of two proteins that have no propensity to associate. In marked contrast E_{P870} was over 20 % in the corresponding RC-LHCII or RC-LHCI chimera. For all chimeras the values of E_{FL} derived from LHC emission data were in agreement with values of E_{P870} derived from RC absorbance data (**Table 5.2**). This correspondence between independently-determined efficiencies from separate data sets reinforced the conclusion that energy transfer was taking place from the plant LHCs to the bacterial RCs within the chimera. The exception to this was a variance between E_{P870} and E_{FL} for samples where WT RCs were mixed with LHCI-Td or Td-LHCI-Td. As discussed above, it is likely that the decrease of emission on adding the RC to either LHCI was due to a change in the latter's quantum yield as there was no accompanying enhancement of P870 photobleaching in the same system. Hence the values of E_{FL} for these two cases are likely unrealistic.

By either method the estimated efficiency of energy transfer in the RC#LHCI#RC chimera was higher than that in LHCI#RC chimera (**Table 5.2**), consistent with the presence of two acceptors in the former, but it was notable that efficiency was not equally increased. Estimates of the energy transfer efficiency to the second RC added to Lhca1 were made using**:**

$$E_{FL,P870} = \frac{E_{a1} + E_{a4} - 2E_{a1}E_{a4}}{1 - E_{a1}E_{a4}} \qquad (\mathbf{5.4})$$

where $E_{FL,P870}$ was the apparent energy transfer efficiency for the RC#LHCI#RC chimera estimated from either LHC fluorescence or P870 photobleaching and E_{a4} was the corresponding single donor/acceptor pair energy transfer efficiency for the LHCI#RC chimera where the RC is attached to Lhca4 and E_{a1} for RC attached to Lhca1. From E_{FL} the value of E_{a1} was 11.5 % (compared to E_{a4} = 19.7 %) and from E_{P870} the value of E_{a1} was 13.9 % (compared to E_{a4} = 20.1 %). Hence this analysis yielded values for transfer to the second RC attached to Lhca1 that were only either 58 % or 69 % of that for transfer to the first RC attached to Lhca4. This is consistent with the presence in the Lhca4 subunit of the relatively low energy red-form chlorophyll *a* dimer (**Fig. 5.1C**) that is responsible for the red-enhancement of the LHCI emission spectrum[74,76,180,240], and which may have produced more efficient energy transfer to an attached RC than was the case for Lhca1.

Also by either estimate, the efficiency of energy transfer in the LHCI#RC chimera was not significantly higher than either the LHCII#RC or RC*4*#dLHCII chimera despite a ~80 %

stronger spectral overlap factor *J* (**Table 5.2**). This is likely due to the reconstituted LHCI heterodimers being in a partially quenched state, reducing the quantum yield to only 29 % of that of LHCII[74]. This was also the case in the present work where the estimates of quantum yield were 30 % for LHCI-Td and 28 % for Td-LHCI-Td (**Table 5.2**). It is plausible that the ~70 % lower quantum yield of LHCI relative to LHCII offset the ~80 % stronger spectral overlap, producing similar energy transfer efficiencies in the LHCI and LHCII chimeras. In future work it might be possible to partially overcome this through modification of LHCI in the native organism, as the quantum yield of native LHCI heterodimers has been reported to be ~64 % that of LHCII, more than double that of recombinant LHCI[74].

5.2.9. Energy transfer efficiencies on an electrode

On an electrode surface, energy transfer was enabled in both mixtures of unadapted proteins and in chimeras. Comparison of the EQE spectrum of the dLHCII + WT RC mix (**Fig. 5.4D**) with those of the LHCII#RC and RC*4*#dLHCII chimeras (**Fig. 5.13A**) revealed that there was little difference between them, meaning that the unbound RC and LHCII proteins in the mixture were within FRET distance of one another to a similar extent as that endowed by linking with SpyTag/SpyCatcher. Estimation of energy transfer efficiencies for protein mixtures and chimeras on electrodes was achieved by quantifying the contribution of the LHCII absorbance to the RC photocurrent in a similar approach to that used to calculate the energy transfer efficiency in solution from the data on RC photochemistry. This involved applying a version of **Eq. 5.3** where k_f was replaced by EQE values and the absorbances of LHCII and RC at the corresponding wavelength. Mean values of the apparent energy transfer E_{app} over the wavelength range 640 nm to 680 nm were calculated and are summarized in **Table 5.3**.

The estimates did not reveal a clear advantage to tethering the two proteins in a chimera before adsorption to the electrode under the experimental conditions used for the photocurrent measurements. This was likely to be due to a high packing density on the electrode surface that promoted energy transfer between adjacent untethered LHCIIs and RCs, as depicted in the schematics in **Fig. 5.15** for two types of system, a mixture of WT RCs and dLHCII and the RC*4*#dLHCII chimera, where the associated values of E_{app} were similar.

5.2.10. Effect of packing density on energy transfer efficiency for mixtures and chimeras

To explore the effect of different surface packing densities on the LHCII/RC energy interaction, a Monte Carlo simulation with Metropolis approach was implemented in MATLAB based on

the polygon functions. The purpose of this simple simulation was to explore any difference brought about by pre-conjugating LHCII and RC in a chimera before adsorbing to an electrode surface. RC and LHCII complexes in DDM micelles were considered to be hard objects that had no propensity for interaction. Paired LHCII and RC geometries were constrained by SpyTag/SpyCatcher at the N-terminus of protein in the case of the chimeras. Because proteins were treated as hard objects no overlap was allowed during the simulation, and therefore no temperature input was considered. Inter-centre distances were extracted from the position of every LHCII relative to the closest RCs at various extents of occupation ranging from 5% to 55%, as described Section 1.18.4. Unit cells are depicted in **Fig. 5.16** for a mixture of unadapted LHCII and RC and **Fig. 5.17** for an RC-LHCII chimera. In the simulation each unit cell was surrounded by 8 identical cells to achieve periodic boundary conditions and the closest center to center distance of LHCII to every RCs were extracted from simulation.

Table 5.3. Apparent energy transfer efficiency on an electrode

System	E_{app} (%)
WT RC + Td-dLHCII	33.1 ± 2.8
RC*4*#dLHCII	37.9 ± 4.9
LHCII#RC	38.9 ± 4.6
LHCI-Td + WT RC	13.9 ± 1.3
LHCI#RC	23.6 ± 1.2
$(WT\ RC)_2$ + Td-LHCI-Td	35.6 ± 1.1
RC#LHCI#RC	29.7 ± 1.6

E_{app} was estimated from EQE action spectra and absorption spectra of samples over the wavelength range from 640 nm to 680 nm to cover the two Chl Q_y bands.

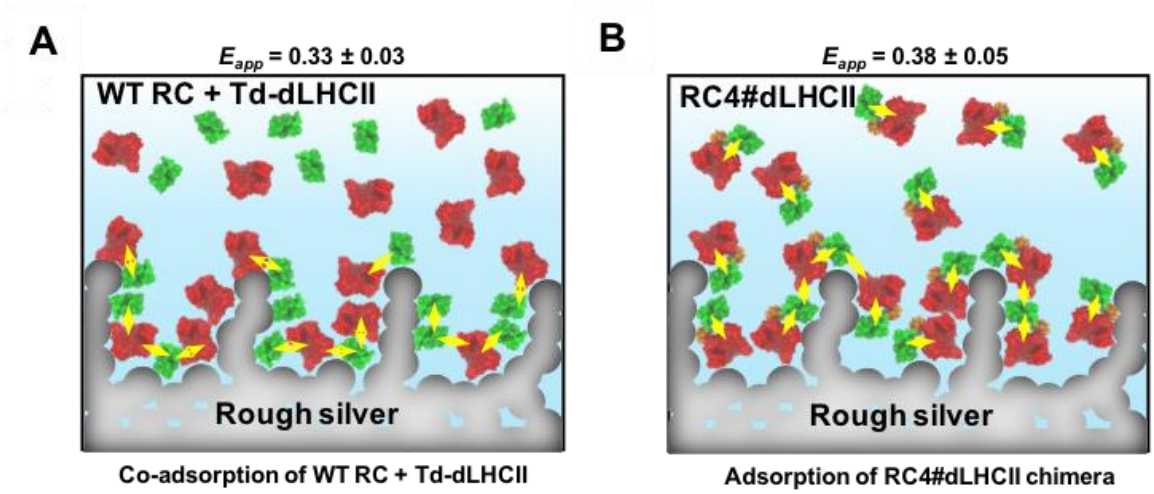


Figure 5.15. Schematics of adsorption of proteins on a roughened silver electrode. (A) Adorption of two independent populations of LHCII and RC to the surface. (B) Adsorption of a SpyTag/SpyCatcher linked LHCII-RC chimera to the surface. Red objects – RC; green objects – LHCII; orange objects – SpyTag/SpyCatcher complex; yellow arrow – energy interaction. Energy transfer efficiencies on the surface are comparable if the packing density is sufficiently high.

The extracted centre-centre distances were then used to estimate the energy transfer efficiency from LHCII to RC based on a FRET mechanism using **Eq. 3.8** (see Chapter 3) to obtain estimations of the single LHCII to single RC FRET efficiency ($E_{DA}(i)$) for each LHCII to all RCs. The apparent energy transfer efficiency E_{app} can then be calculated with **Eq. 5.5**:

$$E_{app} = \sum_{i=1}^{N_{RC}} \frac{E_{DA}(i)}{1 - E_{DA}(i)} \frac{1}{1 + \sum_{i=1}^{N_{RC}} \frac{E_{DA}(i)}{1 - E_{DA}(i)}} \qquad \textbf{(5.5)}$$

where N_{RC} is the number of RCs in each simulation (N_{RC} = 25). In fact, rearrangement **Eq. 5.5** when N_{RC} = 2 yields the expression of **Eq. 5.4**.

The computed E_{app} at all tested surface fractional occupancies is shown in **Fig. 5.18**. A value of κ^2 = 3 was used during the calculation of $E_{DA}(i)$ as this gave a value for E_{app} at very low surface occupation that was near the experimentally-observed efficiency of energy transfer for chimeras in solution (~20%).

The point to note from **Fig. 5.18** was not the absolute values of E_{app} but rather the trends with surface occupancy when LHCII and RC were independent or covalently pre-assembled.

The trends revealed by the simulation reinforced the observation that in a diluted condition, either in solution or on a 2D surface, pre-association of LHCII and RC in a chimera led to observable energy transfer while the lack of any propensity for association prevented any energy communication in a mixture of unadapted proteins. In contrast, at high packing density the co-existence of two populations of protein on a surface with a high concentration resulted in a more similar energy transfer efficiency regardless of whether they were pre-assembled in a chimera or not.

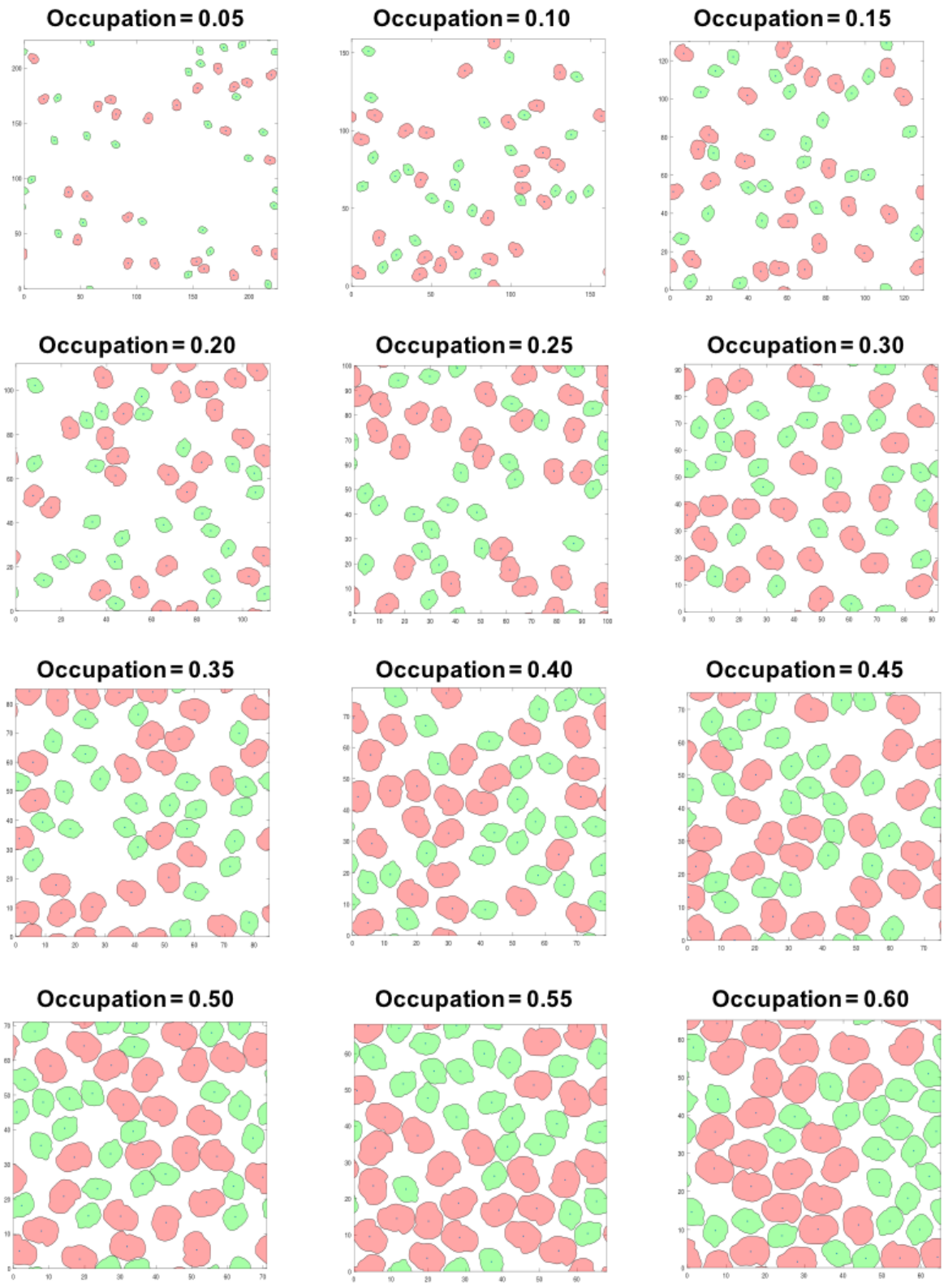


Figure 5.16. Packing of equimolar unadapted LHCII and RC at eleven percentage surface occupations. Green objects represented WT LHCII and red objects represented WT RC. There was no propensity for association between LHCII and RC or with themselves. Dots indicate the centroid of every object used to measure distances.

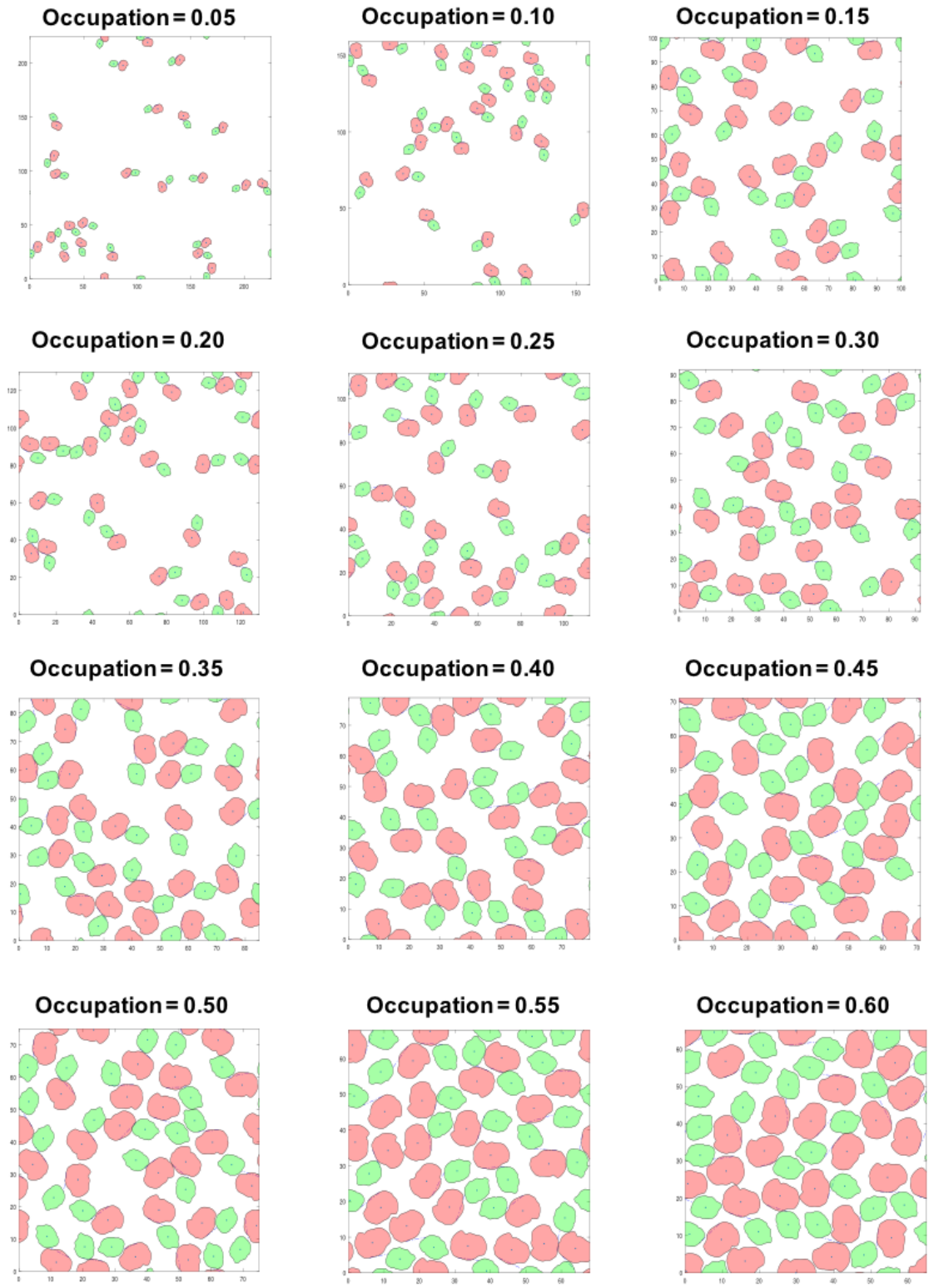


Figure 5.17. Packing of LHCII#RC chimeras at eleven percentage surface occupations. Green objects represented LHCII and red objects represented RCs within each chimera. Each LHCII was constrained with one RC at a range of distances to realize the covalent interaction endowed by SpyTag/SpyCatcher. Dots indicate the centroid of every object used to measure distances.

5.3. Discussion

Adaptation of LHCII or LHCI proteins with SpyTag and RCs with SpyCatcher enabled efficient formation of a pigment-protein chimera that, in contrast to mixtures of one or more unadapted proteins, engaged in energy transfer from the Chl light harvesting system of the LHC to the BChl charge separation system of the RC. This expanded the light harvesting capacity of the RC into regions where its native pigments have very low absorbance, specifically between 620 and 700 nm and between 400 and 500 nm. Chimeras were stable, could be formed in high yield, and could be purified with predictable ratios of LHC and RC absorbance. The efficiency of energy transfer could be assessed from either the quenching of LHCII emission or enhancement of RC photobleaching on chimera formation and was in the region of 20-30% in the systems constructed.

Although, as evidenced by the EQE action spectra, both tethered and untethered LHC-RC systems were capable of energy transfer on a photocathode, tethering within a chimera also conferred a number of advantages under these circumstances. First, it is possible to say with certainty that the stoichiometry of LHCs per RC was 1:1 for electrodes coated with heterodimeric chimeras such as RC4#dLHCII or LHCII#RC, or the RC#LHCI#RC heterotetramer, and was 2:1 for electrodes coated with LHCI#RC heterotrimers. In contrast, because it cannot be assumed that the electrode binding affinity of RCs and LHCIIs is the same, the actual stoichiometry of untethered proteins such as WT RCs and dLHCII complexes on any given electrode is uncertain. In nature the numbers of each type of pigment protein in a photosynthetic membrane is tightly controlled genetically and can be varied in response to environmental conditions such as light intensity.

Second, tethering ensured that each RC was in close proximity to at least one LHC and *vice versa*, whereas uncontrolled absorption of RCs and LHC complexes on a surface leads to the possibility of uneven protein distribution with the formation of domains of either LHCs or RCs as the result of random doping to the surface. This might be the reason for the slightly lower energy transfer efficiency observed in a mixture of unadapted LHCII and RC proteins than in the chimeras (**Table 5.3**), which could be confirmed from a simple hard-body Monte Carlo simulation (**Figs. 5.16** and **5.17,** bottom panels). In nature the genetic encoding of protein-protein interfaces controls the spatial arrangement of different pigment-proteins in the membrane such that the environment of each is optimised for efficient energy harvesting or electron transfer, and tethering provides a means of mimicking this.

Third, tethering a RC to a LHC limits freedom in their mutual orientation, reducing heterogeneity in the distances and orientations between their Chl and BChl pigment systems.

Both of these factors are important for the efficiency of FRET and are controlled in nature by genetically-encoding uniformity in the orientation of each pigment-protein complex in the two-dimensional membrane system, and genetically-encoding protein-protein interactions such that each pigment-protein is positioned adjacent to other specific pigment-proteins in highly-defined architectures. Such precision in two- or three-dimensional architecture is impossible to achieve through drop-casting of a mixture of disparate proteins onto a uniform substrate, but the type of modification described above illustrates how genetic modification can be used to begin to gain control over how hybrid photosystems self-assemble by fixing the association of, in this case, a Chl light harvesting system with a BChl charge separation system.

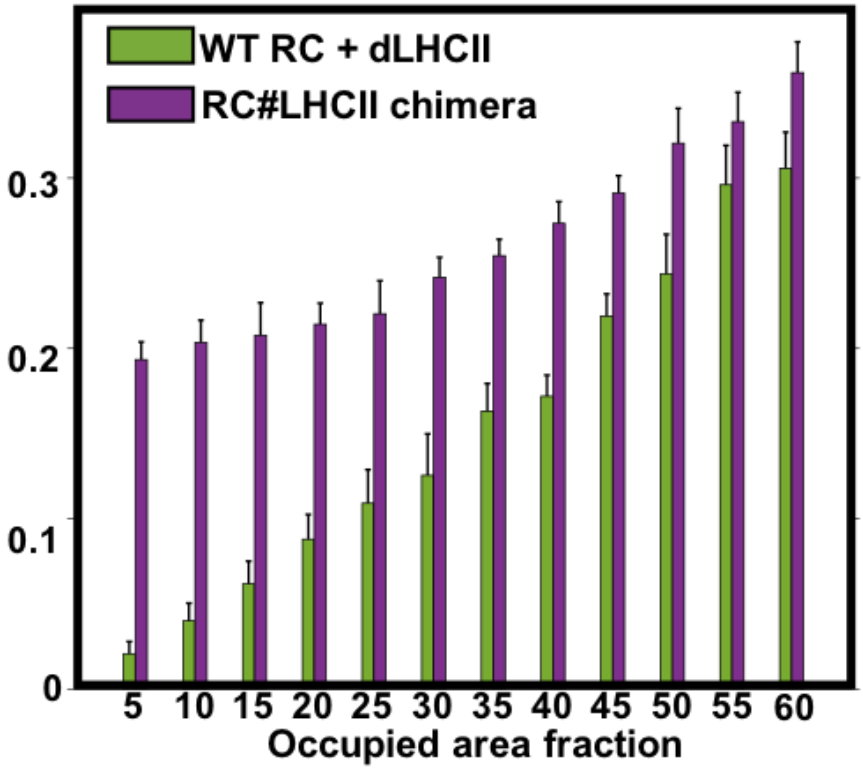


Figure 5.18. Estimated energy transfer efficiencies for WT mixtures and RC#LHCII chimera. Theoretical calculation of apparent energy transfer efficiency from ten repeat simulations of packing a WT RC + dLHCII mixture or the same amount of LHCII#RC chimera on an open boundary 2D surface.

Although, at high surface coverages, it turned out that tethering did not confer a large advantage in terms of the efficiency of LHC to RC energy transfer, it should be remembered that not all biohybrid devices for solar energy conversion require protein deposition on the photoelectrode. In many devices it has been demonstrated that it is possible to generate photocurrents or photovoltages from RC proteins in solution, and in the future development of these devices it would be expected that building disparate pigment-proteins into specific, functional architectures through genetic modification should be advantageous over just mixing proteins that have complementary photophysical properties but have no natural ability to engage in productive interactions such as those seen in complex natural photosystems. Tethering would also bring advantages if it were required to deposit proteins on a surface at low densities, as illustrated in the simulation.

It is interesting that for chimeras involving attachment of RCs at the C-terminus of LHCII the length of the linker connecting the RC to the SpyCatcher/Tag module seemed to affect the amount of quenching of LHC emission observed, a stronger effect being seen with no or four amino acid linker than with a eight amino acid linker (**Fig. 5.4A**, with LHCII-T). In the absence of structural information it is not possible to be certain over the interpretation of this observation, which also needs verification through a more extensive and systematic analysis, but the simplest preliminary interpretation would be that the shorter linkers produce more compact structures that are more favourable for energy transfer due to reduced donor-acceptor distances.

5.4. Conclusions

This work shows that genetically adapting two types of evolutionarily-disparate photosynthetic membrane proteins with an extramembrane interface domain enables self-assembly of a chimeric photosystem in which UV/near-IR solar energy conversion by a BChl RC is augmented by visible light capture by Chl LHCs. This approach inspired by a concept of synthetic biology, adapting naturally incompatible components to interface in a predictable way through genetic encoding, creates covalently-stabilised macromolecular photosystems that are predictable and programmable. In addition to providing novel photosynthetic structures and energy transfer pathways to explore, these polychromatic photosystems constitute interesting new materials for biohybrid devices that in recent years have expanded in application beyond photoelectrochemical solar energy conversion to fuel molecule synthesis, energy storage, biosensing, touch sensing and photodetection. Finally, the demonstrated flexibility with which RCs and LHCs could be interfaced opens the possibility of constructing more elaborate, self-assembling chimeric photosystems that employ multiple

orthogonal linking modules[157,242] and a wider range of photosynthetic and redox proteins that, despite being separated by billions of years of evolution, can be adapted for future solar energy conversion through genetic programming of predictable interfaces.

5.5. Supplementary figures

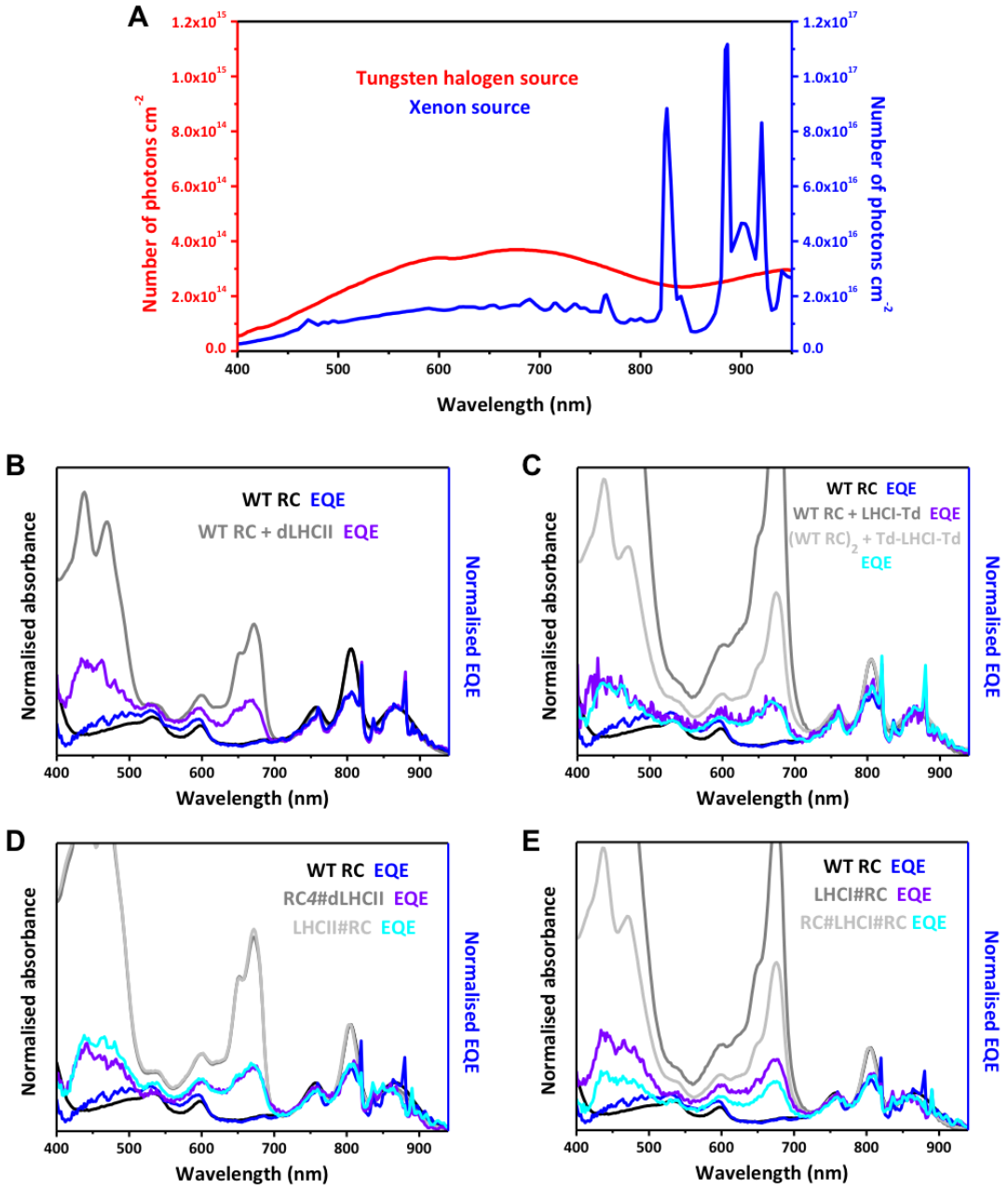


Figure S5.1. Lamp spectra and EQE spectra recorded with a xenon lamp. (A) Output spectra for tungsten and xenon lamps used to record EQE action spectra. (B) Solution absorbance and EQE spectra for WT RCs compared with those for 1:1 mixtures of WT RCs and dLHCII. (C) Solution absorbance and EQE spectra for WT RCs compared with those for 1:1 mixtures of WT RCs and LHCI-Td or Td-LHCI-Td. (D) Solution absorbance and EQE spectra for WT RCs compared with those for RC*4*#dLHCII or LHCII#RC chimeras. (E) Solution absorbance and EQE spectra for WT RCs compared with those for LHCI#RC or RC#LHCI#RC chimeras. The EQE action spectra were measured with a xenon lamp. In (B)-(E) the

absorbance spectra were normalised at 804 nm, whilst each EQE spectrum was normalised to the corresponding absorbance spectrum at the maximum of the P870 band.

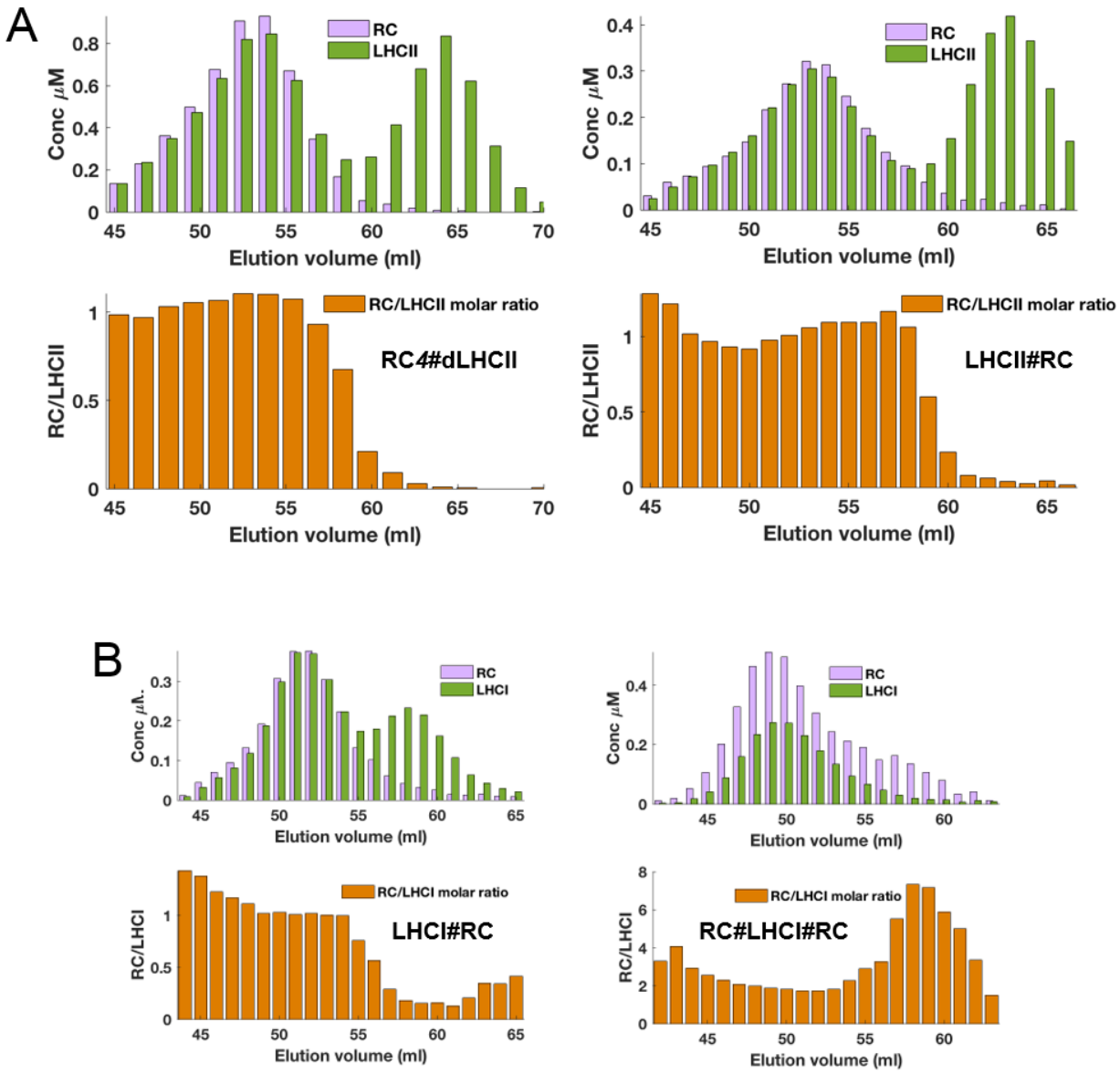


Figure S5.2. Purification of covalent RC-LHCII chimeras. (A) Purification of chimeras by gel filtration chromatography for RC*4*#dLHCII (left) and LHCII#RC (right). Plots show concentrations of RC and LHCII in fractions derived from absorbance spectra, and the resulting RC/LHCII ratio was used to identify fractions predominant in the desired 1:1 RC/LHCII chimera product. (B) Equivalent data for the LHCI#RC and RC#LHCI#RC chimeras where the RC:LHC ratio were 1 and 2, respectively.

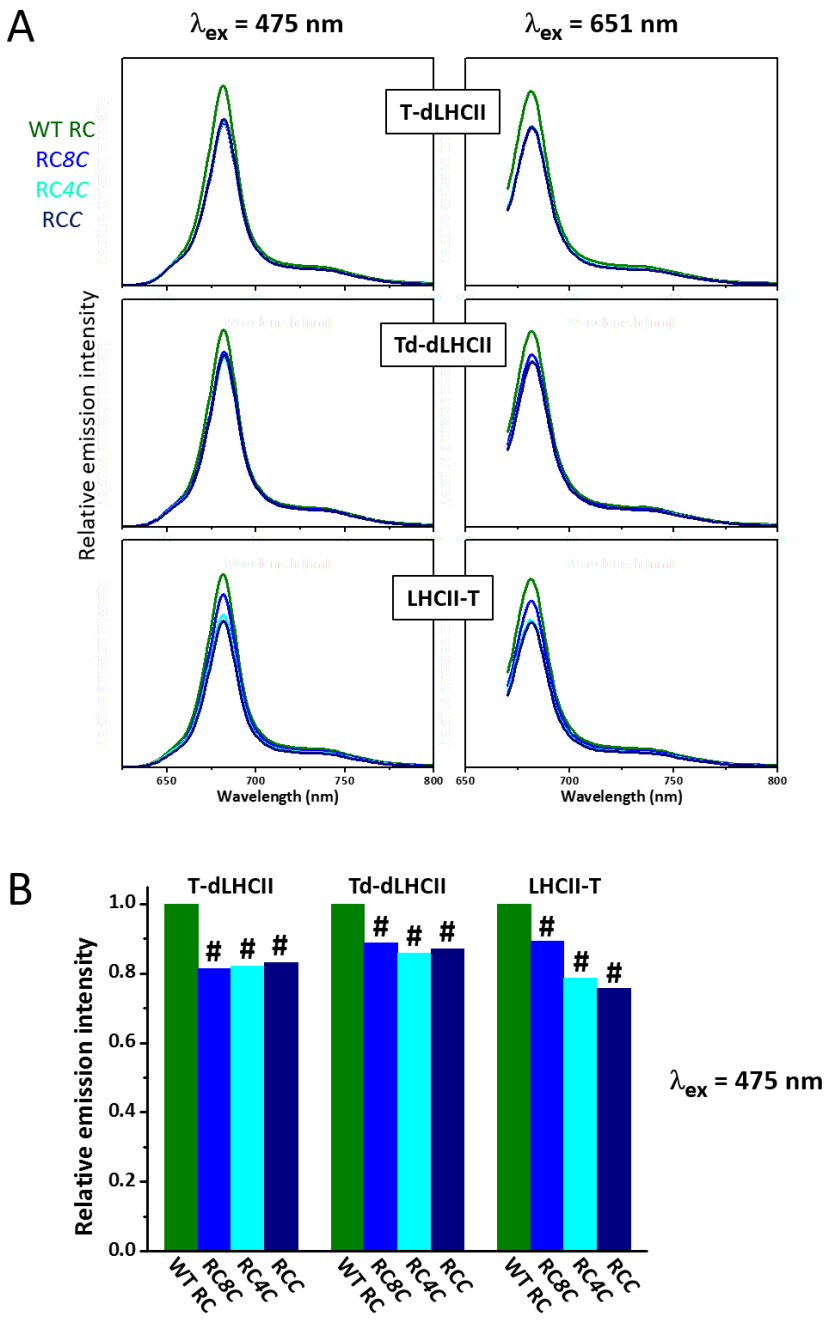


Figure S5.3. Emission quenching in RC-LHCII chimeras. (A) Spectra recorded with two excitation wavelengths for mixtures with a 2:1 RC:LHCII composition (to ensure 100% incorporation of LHCII into chimeras) demonstrating that conjugation with SpyCatcher-adapted RCs quenches emission from all three SpyTag-adapted LHCII complexes (shades of blue) relative to that measured in the presence of unadapted WT RCs (green). (B) LHCII emission at 681 nm in RC-LHCII heterodimers (shades of blue) relative to that in a concentration-matched mixture of the component adapted LHCII and WT RCs (green), following excitation at 475 nm. Symbol # denotes mixtures that self-assemble into chimeras. Equivalent data obtained with 651 nm excitation are shown in **Fig. 5.11A**.

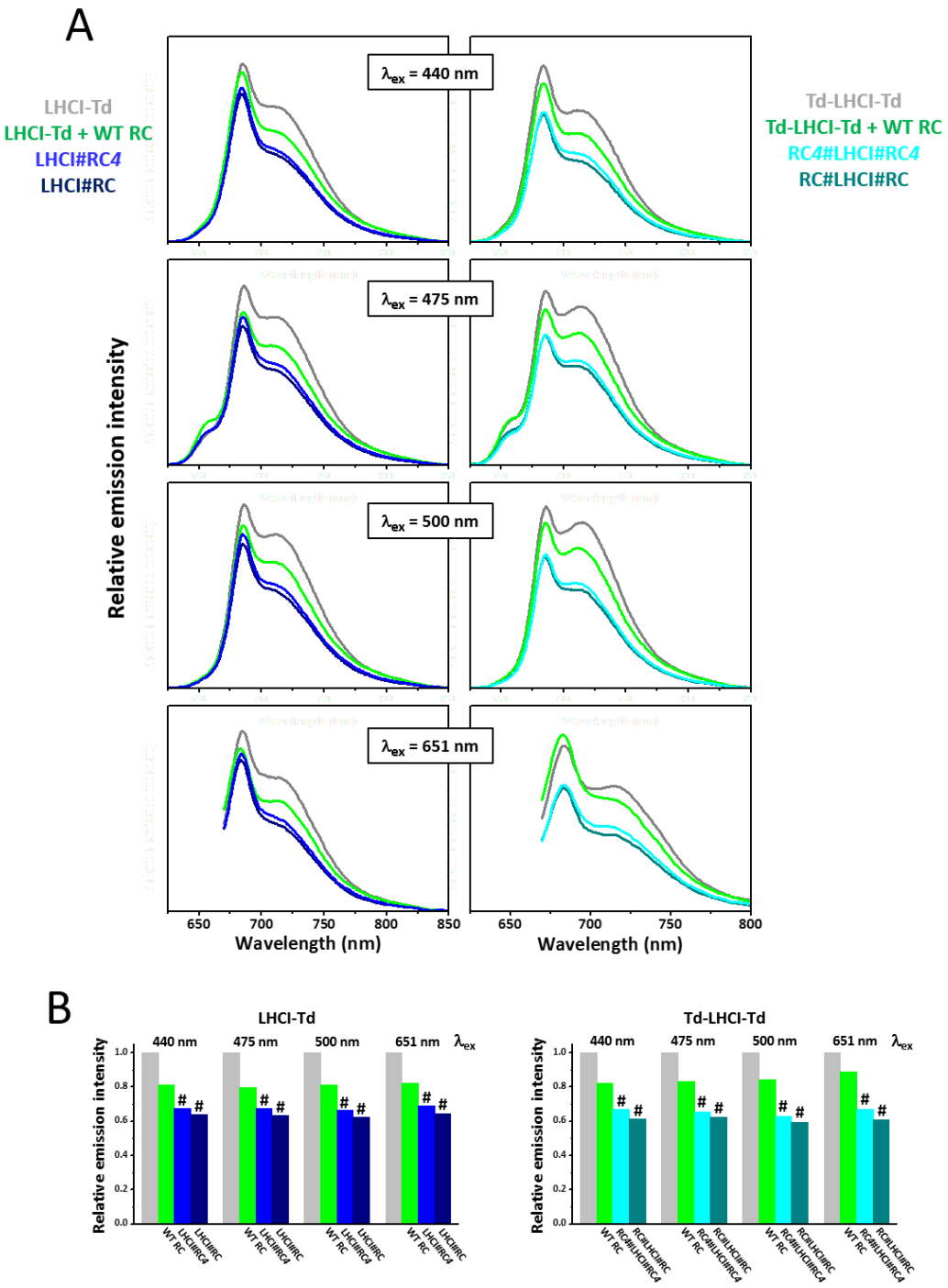


Figure S5.4. Emission quenching in RC-LHCI chimeras. (A) Data are shown for four excitation wavelengths, with data for RC-LHCI chimeras on the left and $(RC)_2$-LHCI chimeras on the right. Conjugation with SpyCatcher-adapted RC*C* or RC*4C* quenches emission from LHCI-Td (left) or Td-LHCI-Td (right) relative to that measured for LHCI in the presence of an equivalent concentration of WT RCs (green) or for LHCI alone (grey). (B) LHCI emission at 714 nm in RC-LHCI chimeras (left) or $(RC)_2$-LHCI chimeras (right) relative to that in a concentration-matched mixture of the component adapted LHCI complex and WT RCs (green)

or LHCI complexes alone (grey), following excitation at four different wavelengths. Symbol # denotes mixtures that self-assemble into chimeras. The data collected with 651 nm excitation are also shown in **Fig. 5.12A**.

Table S5.1. Fits of P870 photobleaching and photon absorbance by RCs and LHCs.

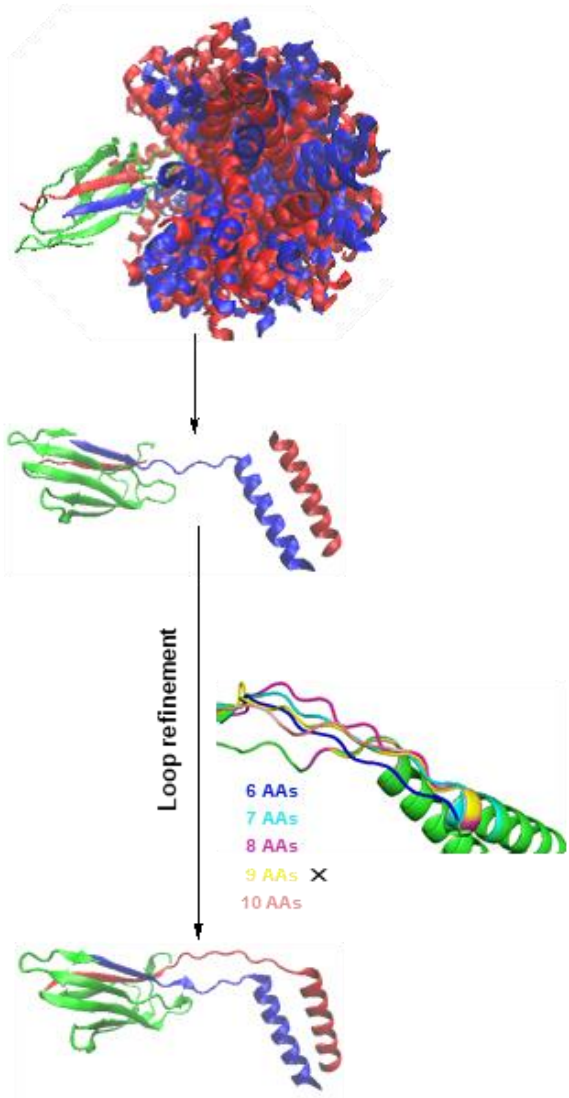
Sample	k_f [a] (10^{-2} s^{-1})	k_r [a] (10^{-2} s^{-1})	RC photon absorbance[b] (10^{12} $cm^{-2}s^{-1}$)	LHC photon absorbance[b] (10^{12} $cm^{-2}s^{-1}$)
WT RC	6.2 ± 0.8	79.3 ± 11.7	2.04	0.0
RC*C*	2.6 ± 0.5	59.5 ± 15.7	2.34	0.0
RC*4C*	2.4 ± 0.5	56.9 ± 15.5	2.26	0.0
WT RC + dLHCII	6.9 ± 0.7	72.3 ± 8.9	2.12	28.8
RC*C* + dLHCII	2.5 ± 0.4	46.1 ± 11.2	1.93	26.2
RC*4C* + dLHCII	2.7 ± 0.4	51.4 ± 10.3	2.49	26.6
WT RC + Td-dLHCII	7.9 ± 0.5	44.5 ± 3.7	2.04	29.3
RC*4*#dLHCII	10.1 ± 1.1	66.3 ± 8.5	1.55	22.9
WT RC + LHCII-T	7.2 ± 0.7	67.0 ± 7.7	1.95	27.7
LHCII#RC	11.3 ± 0.9	71.2 ± 7.0	1.93	27.9
WT RC + LHCI-Td	8.0 ±1.3	58.8± 12.1	1.00	23.0
LHCI#RC	14.3 ± 1.5	73.8 ± 9.2	1.29	28.9
$(WT\ RC)_2$ + Td-LHCI-Td	8.5 ± 0.9	79.9 ± 10.2	1.97	22.2
RC#LHCI#RC	11.1 ± 0.7	70.4 ± 5.1	3.40	38.2

[a] Fitted first order kinetic constants ± standard error.

[b] Number of photons absorbed by RCs and LHCs per unit area per second calculated from integration across spectra of the product of incident light power and protein absorbance (1 - transmission).

Chapter 6

Evaluation of Alternative Interface Components for Photosystem Construction



Abstract

A self-organizing photosynthetic membrane is fundamental to the high efficiency of the energy harvesting function of photosynthesis. An aim of this work was to mimic natural strategies for organisation in synthetic photosystems using a range of linking modules provided by recent developments in synthetic biology. In the previous chapter use of the SpyTag/SpyCatcher system was explored. In this chapter *de novo* designed coiled coils, a split intein and a newly designed “SpyCoil” hybrid system were also evaluated for the purpose of the bottom-up redesign of photosynthesis. Coiled coil and split intein linking strategies showed positive results in enabling self-assembly of genetically engineered LHCII and the purple bacterial RC into a single module but the efficiency of desired product formation was low. The SpyCoil hybrid system showed positive results when tested with eGFP and a combination on eGFP and LHCII. The principal conclusion was that further improvements are needed in the efficiency and construability of these linking modules in the context of bottom-up redesign of photosystems.

6.1. Introduction

The previous chapter described an effective approach to the bottom-up construction of a self-assembling all-biological chimeric photosystem comprising purple bacterial RCs and plant LHCII or LHCI with designed oligomeric states. To fully realize the complexity of natural photosystems, effective alternative means to build assemblies with more than two components are also desirable. Advances in synthetic biology in recent years have offered a range of potential peptidyl linking strategies besides the SpyTag/SpyCatcher[147] system that could be used to either add complexity or as a substitute for SpyTag/SpyCatcher. Two of these, *de novo* designed coiled coils[146] and split inteins[166] have been described previously in the literature. The third, a hybrid system dubbed "SpyCoil" [148,149] was developed and evaluated as part of this project. All three linking strategies are ways to achieve peptidyl-based self-assembly that can be realized *in vitro* and *in vivo* through genetic modification.

Coiled coils are formed by α-helices that, as a consequence of their sequence, can self-assemble into higher-order structures with various oligomeric states, orientations and combinations[146]. They are based on a heptad repeat of hydrophobic (*H*) and polar (*P*) amino acids with the sequence *HPPHPPP*, the identity of the amino acid at each position controlling the oligomeric state of the coiled-coil bundle and parallel/anti-parallel topology. They are now one of the most understood protein structures, whose specificity and affinity can be tuned accurately for desired properties[159]. For example, in the heterodimeric coiled coil using in this project, the CCA/CCB pair, the two helices carry oppositely charged amino acids (Asp for CCA; Lys for CCB) that enable heterodimeric packing[162]. The linking strategy is therefore to adapt one photosynthetic protein with CCA and a second with CCB, and mix the two to self-assemble a photosynthetic protein heterodimer driven by formation of the coiled-coil heterodimeric bundle. A potential advantage of this strategy is that the linking interaction is non-covalent, and so could be switchable using, for example, a short peptide with a stronger binding affinity for one half of the CCA/CCB coiled-coil bundle to control the energy flow.

The split intein system was developed from self-excising intein protein domains that are able to remove themselves from a polypeptide sequence with formation of a peptide bond between the flanking extein domains. Splitting an intein into two sub-domains leads to the creation of a "clean" linking system in which the N-terminal intein domain on one host protein and the C-terminal domain on a second host protein will combine and self-cleave to leave a short linking "scar" that consists of three amino acids. In the case of Cfa split intein,[166] which was used in this work, this scar has the sequence -CFN-. Although in principle this is a very effective way of covalently linking two proteins using a component that is (almost entirely) not

retained in the final product, intein splicing can be affected by side reactions leading to a low efficiency of formation of the desired product.

The third linking strategy was a hybrid of the coiled coil and SpyTag/SpyCatcher systems in which a heterodimeric coiled-coil bundle is stabilised by a covalent interaction introduced through a variant of the SpyTag/SpyCatcher system. This variant, known as SpyTag/KTag/SpyLigase was developed by further splitting the SpyCatcher domain into two parts dubbed “KTag”, a short 10-residue sequence that includes the reactive lysine and SpyLigase that comprises the rest of the SpyCatcher sequnce[149]. The concept is that SpyTag and KTag can each be engineered onto a protein of interest which, upon mixing with each other and the SpyLigase domain, will be covalently linked through isopeptide bond formation between SpyTag and KTag. This has the potential advantage that the resulting linking domain is much smaller than achieved with SpyTag/SpyCatcher, but has the disadvantage that a third protein reagent, SpyLigase, needs to be added to the SpyTag- and KTag-modified proteins to achieve self-assembly. In addition, due to the additional splitting of the parent CnaB2 domain, the tri-component SpyLigase system is only functional in the presence of a high concentration of trimethylamine *N*-oxide (TMAO), and isopeptide bond formation is significantly less efficient than that achievable through SpyTag/SpyCatcher.

The concept of a hybrid “SpyCoil” system developed from a postulate that efficiency of isopeptide bond formation might be restored to the SpyLigase system if one can provide additional affinity by pre-assembling part of the complex by bringing SpyTag and KTag together. Moreover, because the activity depends on the extra affinity provided by the coiled-coil binding motif, the specificity could be readily tuned by using designed binding domains.

The work described in this chapter evaluated use of heterodimeric coiled coils, a split intein and the new SpyCoil hybrid to self-assemble protein heterodimers. It was carried out in parallel with early studies on the formation of RC-LHCII chimeras using SpyTag/SpyCatcher, which emerged during the project as the system of choice in terms of simplicity and efficiency of product formation. To test the coiled-coil approach, 3.5- and 4-heptad repeat versions of the *de novo* designed heterodimeric coiled coils CCA and CCB[162] were fused to both the RC and LHCII, and mixtures evaluated for heterodimer formation. To test the split intein approach, the synthetic Cfa system was used which was derived as a consensus sequence from an alignment of 73 natural DnaE intein sequences that display fast splicing[10]. The Cfa^N large domain was fused to the RC and the smaller Cfa^C domain was fused to LHCII. To test the SpyCoil approach, the *de novo* designed heterodimeric coiled coils P3 (basic) and P4 (acidic)[148] were fused to SpyTag and KTag, respectively This created SpyCoil components

that successfully restored the activity of the SpyTag/KTag/SpyLigase reaction under normal buffer conditions. The system was tested on eGFP and was successfully used to attach eGFP to LHCII.

6.2. Results

6.2.1. Evaluation of the use of coiled coils to form arrays of RCs

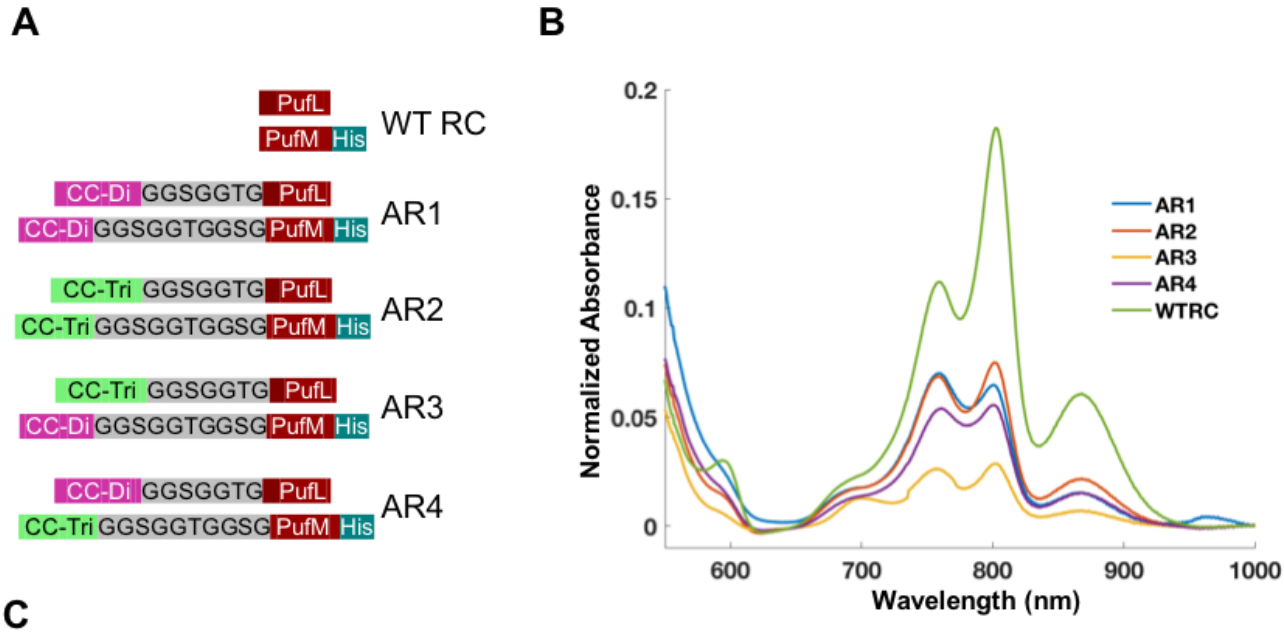
The first step in evaluating the use of coiled coils for protein linkage was to extend previous work by David Swainsbury and others on the formation of RC homo-oligomers and hetero-oligomers using coiled coils[243]. This work demonstrated assembly of monomeric RCs into dimeric, trimeric and tetrameric forms through the addition to the N-terminus of either PufL or the symmetrical PufM of a four-heptad repeat coiled-coil sequence[244] encoding a homodimer (CC-Di), homotrimer (CC-Tri) or homotetramer (CC-Tet). The work also demonstrated the formation of RC heterodimers through modification of the N-terminus of PufL with a 3.5-heptad repeat acidic (CCA) or basic (CCB) coiled-coil sequence[146].

This work was moved forward by examining whether RCs are able to self-assemble into a larger network if both PufL and PufM are modified with sequences for homo-type coiled-coil bundles. This presents a challenge for the assembly of individual RCs, as two of the three subunits that make up the protein have added sequences that code for dimeric or trimeric coiled coils, and there is the additional potential complication of a RC network or mesh disturbing the structural integrity of RCs once assembled, or the general membrane architecture.

6.2.1.1. Engineering and purification of double coiled-coil RCs (AR1 - AR4)

RCs were made in the combinations L-Di/M-Di (AR1), L-Tri/M-Tri (AR2), L-Tri/M-Di (AR3) and L-Di/M-Tri (AR4), where L-Di denotes a homodimer coiled-coil sequence (CC-Di) fused to the N-terminus of PufL and so on (**Fig. 6.1A, Table S6.1**). Remarkably, RCs were expressed in all four strains. An assessment of the expression levels of the four AR mutants based on normalised and scatter-corrected absorbance spectra of photosynthetic membranes are shown in **Fig. 6.1B.** The high ratio of A_{760}/A_{803} indicated presence of some free BChl in these membranes, which implied assembly of RC monomers is affected by the modifications, but the bands at 803 and 865 nm clearly showed the presence of properly assembled RCs. The level of RC expression in the membrane depended on the particular combination of coiled-coil modifications and ranged from ~40 % to ~15 %.

The AR mutant RCs were purified by nickel affinity chromatography and gel filtration (Section 2.5.1). His-tagged WT RCs typically bind to a Ni-column with a 40-50 % yield, and this was also seen for the AR mutants (**Fig. 6.1C)** and so the final yield of RCs was determined mainly by the expression level. The final purity of the AR mutants was not as high as for the WT RC, with A_{280}/A_{803} ratios being high for the AR1 and AR3 mutants in particular (**Fig. 6.1BC**). This is typical of what is seen for RC mutants with lowered expression levels.



C

Sample	Nickel binding%	A_{280}/A_{803}
AR1	45.3	3.12
AR2	47.8	1.69
AR3	34.5	2.18
AR4	40.6	1.44
WTRC	47.0	1.24

Figure 6.1. Constructs and quality of array RC mutants. (A) AR mutant constructs with the adducts highlighted. More detailed sequence information can be found in **Table S6.1.** (B) Membrane absorption spectra of the WT RC and AR mutants. The spectra were corrected for membrane concentration at 1000 nm and then light scatter. (C) Recovery of AR mutants after nickel affinity chromatography and the final product purities A_{280}/A_{803} after gel-filtration chromatography.

6.2.1.2. AR mutants show a low level of oligomerization at high salt concentration

Evaluation of purified AR3 (L-Tri/M-Di) and AR4 (L-Di/M-Tri) RCs by sucrose gradient ultracentrifugation showed that each existed as mixture of dimer and trimer **(Fig. 6.2A)**. As with RC homodimers (L-Di) and homotrimers (L-Tri), both AR3 and AR4 migrated further than the WT RC monomer. However, unlike the singly modified L-Di and L-Tri RCs, the doubly-modified AR3 and AR4 RCs migrated as more diffuse bands pointing towards greater compositional heterogeneity. This interpretation of the images of intact gradients was supported by spectroscopic analysis of fractionated gradients that showed AR3 and AR4 comprised a mixture of dimers and trimers (**Fig. 6.2B**).

This analysis of the oligomeric state formed by the AR3 and AR4 RCs was carried out under standard buffer conditions of 20 mM Tris (pH 8.0) supplemented with 200 mM NaCl and 0.1 % LDAO. It was speculated that the failure to detect higher oligomeric states of these RCs may have been the result of buffer conditions, as the coiled-coil sequences used were developed by Fletcher *et al.* using buffer conditions that were different in terms of pH and ionic strength (137 mM NaCl, 10 mM phosphate buffer, pH 7.4)[244]. To examine whether these conditions limited the range of oligomeric states that could be detected, screening was carried out by varying buffer substances, buffer pH and ionic strength. It can be seen from the comparison of migration of WT and AR4 RCs on sucrose density gradients (**Fig. 6.3**) that AR4 remained as mixture of dimer and trimer and showed no change between pH 7 and pH 8 (in 50 mM phosphate/200 mM NaCl/0.1% LDAO). However, the oligomeric state was found to be sensitive to ionic strength, as described in the next section.

6.2.1.3. Inducement of aggregation of the AR mutants by lowering ionic strength

The rapid formation of aggregates of AR3 and AR4 could be triggered by lowering the salt concentration from the standard 200 mM NaCl. In 20 mM Tricine, ≤10 mM NaCl, 0.1% LDAO buffer, pH 7.0 (ionic strength ≤ 0.01M), AR3 and AR4 precipitated in less than one minute on transfer to low ionic strength buffer, while WT RCs were stable in solution under these conditions for least a day. This could be seen on sucrose gradients (**Fig. 6.4**), with the effect of salt concentration seeming to be positively correlated to the degree of protein aggregation (loss of upper band and appearance of an aggregate band at the bottom of the gradient. It remained elusive whether this fast aggregation was caused by coiled-coil interactions producing higher order oligomers beyond dimers or trimers, perhaps as a result of columbic interactions being stronger due to a lower shielding effect from salt ions, or was induced by nonspecific interactions between the RC and the coiled-coil helices, since it was observed that

single coiled-coil RC mutants also slowly precipitated out in a few hours at room temperature[243].

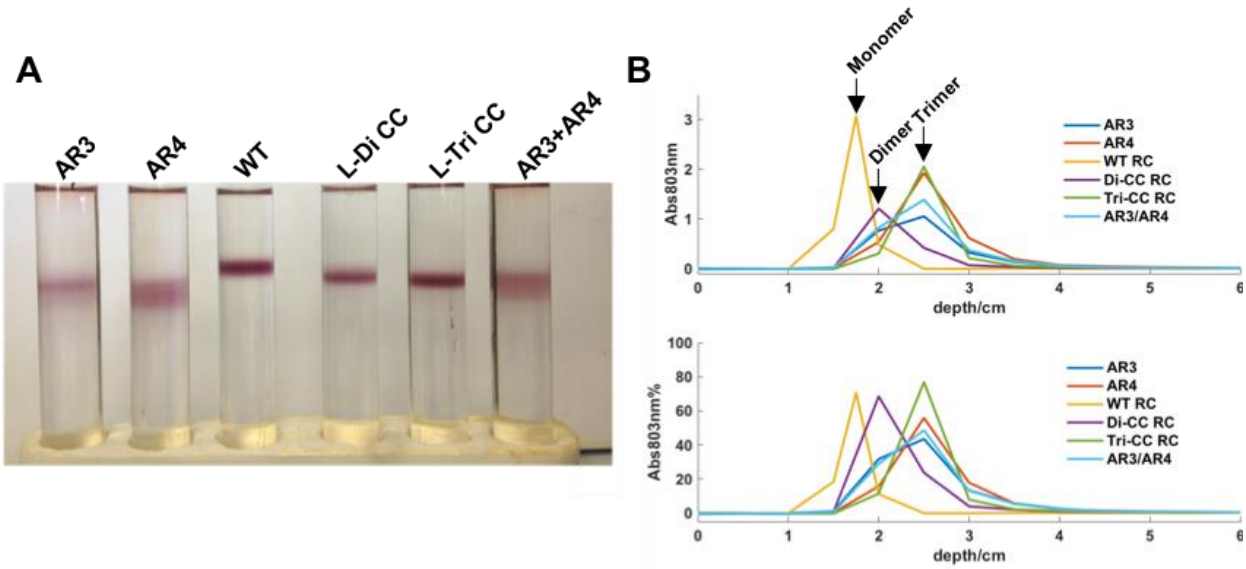


Figure 6.2. Assessment of aggregation states of AR RCs by sucrose density gradient ultracentrifugation. (A) Image of sucrose density gradients. The purple-red band indicated the location of RCs. (B) Direct absorption scanning of the gradient at 0.5 cm intervals using an optical fibre spectrophotometer. The upper panel presents the absolute A_{803} (i.e. RC concentration) after baseline correction. The lower panel indicated the percentage of total RC protein at each position. AR3/AR4 was a 1:1 mixture. Depth is the distance to the top of the gradient. Linear gradients were formed by freeze/thaw of 21 % (w/v) sucrose in 20mM Tris, 0.1% LDAO (pH 8.0)

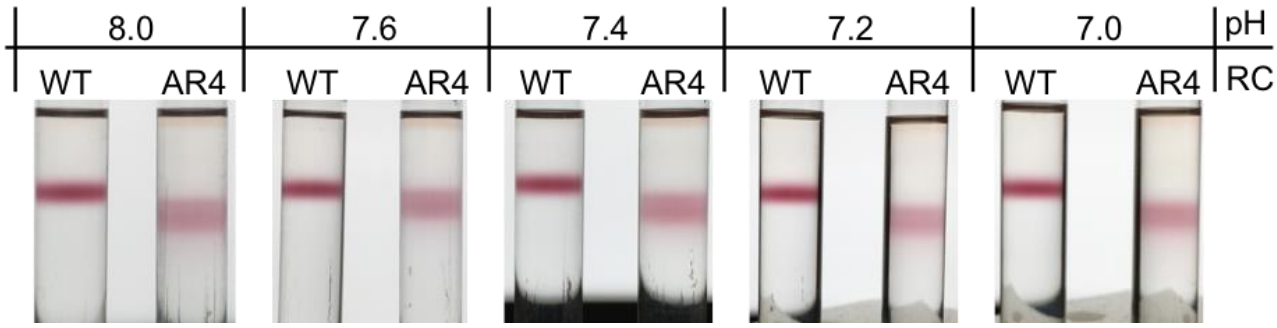


Figure 6.3. Response of aggregation state of mutant AR4 RCs to pH. WT – wild type RC; AR4 – double coiled-coil mutant RC. Linear gradients were formed by freeze/thaw of 21 % (w/v) sucrose in 20mM Tris, 0.1% LDAO (pH 8.0).

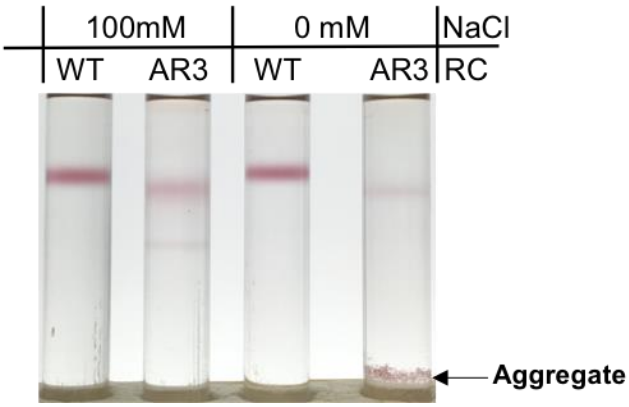


Figure 6.4. Response of AR3 RC to buffer ionic strength. The concentration of NaCl was as indicated. The other buffer components were 50 mM Tricine (pH 7.0)/0.1% LDAO. Linear gradients were formed by freeze/thaw of 21 % (w/v) sucrose in 20mM Tris, 0.1% LDAO (pH 8.0).

6.2.1.4. Array mutant RCs - summary

To summarise, individual RCs were still able to assemble despite double-tagging the PufL and PufM subunits with extramembrane sequences for the formation of homodimeric or homotrimeric coiled coils. Isolated complexes showed compositional heterogeneity, with evidence of larger aggregated states in addition to homodimers and homotrimers. This work, carried out towards the start of the project, gave valuable insights into the importance of buffer conditions for controlling oligomer formation, as well as experience of the techniques used for the assessment of membrane protein oligomers.

6.2.1.5. Evaluation of heterodimeric coiled coils for formation of RC/LHCII heterodimers

Turning to the main theme of this part of the project, the ability of heterodimeric coiled coils to induce dimerization between different photoproteins was evaluated using 3.5- and 4-heptad repeat versions of the acidic CCA and basic CCB coiled-coil sequences developed by Thomas *et al.*[146] that were successfully used in previous work to assemble RC heterodimers[244]. All four coiled coils were added to both the RC and LHCII, producing eight adapted photoproteins.

6.2.2. Design of heterodimeric coiled coil RC and LHCII constructs

For the RC, the four CCA and CCB coiled coils were added to the N-terminus of the PufL by a seven amino acid flexible linker (**Fig. 6.5A, Table S6.1**). The names of the resulting RCs denoted the number of heptad repeats (3.5 or 4), and whether the charged amino acid was lysine (in CCB) or glutamic acid (in CCA). In the 3.5KCC and 3.5ECC RCs the His-tag (His_{10}) was attached to the C-terminus of the M-subunit as for the WT RC (**Fig. 6.5A**). For the 4KCC and 4ECC RCs the His-tag was placed adjacent to the coiled coil; this change was made in response to unpublished observations that with increased size of the added coiled coil, there was a greater tendency for it to be lost through protease activity. Placing the His-tag before the coiled-coil ensured it was retained in the purified protein. The absorbance spectrum of the RC was not affected by any of the coiled-coil modifications (**Fig. 6.5C**). In some Figures, and in the text below, the labels 3.5KCC, 3.5ECC, 4KCC and 4ECC are abbreviated to 3.5K, 3.5E, 4K and 4E.

For LHCII, the same four CCA and CCB coiled coils plus the seven amino acid flexible linker were added to the N-terminally truncated version of Lhcb1.3 described in previous Chapters (**Fig. 6.5B**). The rationale behind removal of the N-terminal 13 residues (3 – 15) was to eradicate any potential influence of this positively charged patch on the acid-base

interaction that gives specificity to the heterodimeric coiled-coil interaction and to gain more control of the distance between the coiled coil and the body of LHCII because this region is also disordered[245]. No spectral alteration was observed when the N-terminal region was replaced with any of the coiled-coil sequences, which indicated the preservation of the function of LHCII as light harvesting complex (**Fig. 6.5D**).

6.2.2.1. Expression and purification of the CCA/CCB RCs and LHCIIs

The four CCA/CCB mutant RCs were purified by nickel affinity chromatography and gel filtration, during which they were detergent -exchanged into Tris/DDM to be compatible with the buffer system used for LHCII (Section 2.5.1). In addition, because the E or K coiled coil carried a strong negative or positive charge, a final step of anion or cation ion-exchange chromatography was applied to eliminate any "damaged" proteins that had lost their coiled-coil modification. Analysis by SDS-PAGE/WB confirmed a good protein quality for all four mutants (**Fig. 6.5E**). The H subunits of all the mutants were well aligned to that of the WT RC and the L subunits were located at higher position than in the WT RC indicating the presence of the additional coiled coil (**Fig. 6.5E**, red arrows). The position of the M subunit band was lower for the 4E/K mutants than the remainder (**Fig. 6.5E**, blue arrows) due to the absence of the His-tag. As expected, anti-His-tag signals were detected for the M subunit in the case of the WT RC and the 3.5E/K mutants and for the higher running L subunit in the case of the 4E/K mutants (**Fig. 6.5E**, marked with H_{10}). For all four mutants SDS-PAGE failed to show a detectable band close corresponding to the WT L subunit, indicating good protein integrity after ion-exchange chromatography (**Fig. 6.5E**). Calculated A_{280}/A_{803} ratios for the mutant RCs were 1.29, 1,27, 1.26 and 1.23 (**Fig. 6.5C**), indicating a high protein purity comparable to that of the WT RC.

After refolding, the four CCA/CCB mutant LHCIIs were also purified by nickel affinity chromatography and gel filtration (Section 2.6.3), followed by anion or cation ion-exchange chromatography. Successful refolding was also confirmed by SDS-PAGE/WB (**Fig. 6.5F**). SDS-PAGE indicated that the purified mutants had only minor contaminants (< 5% according to CS intensity) lacking the coiled coil. The mutants showed apparent changes in mass of the Lhcb1.3 polypeptide that were larger for the 4/3.5E mutants than the 4/3.5K mutants (**Fig. 6.5F**).

The absorbance spectra of the reconstituted LHCII mutants showed small variations relative to one other (**Fig. 6.5D**). This was likely due to variable levels of a loosely bound Chl *b* that contributes to the absorbance around 651 nm (Chl *b* Q_y band) and which can vary

depending on the precise pigment to apoprotein ratio during protein reconstitution. Despite these variances in occupation of this Chl *b* site, the emission spectra of the four coiled-coil mutants were invariant with excitation wavelength at 440 nm, 475 nm and 500 nm (**Fig. S6.4**), diagnostic of correct pigment coordination in functional LHCIIs.

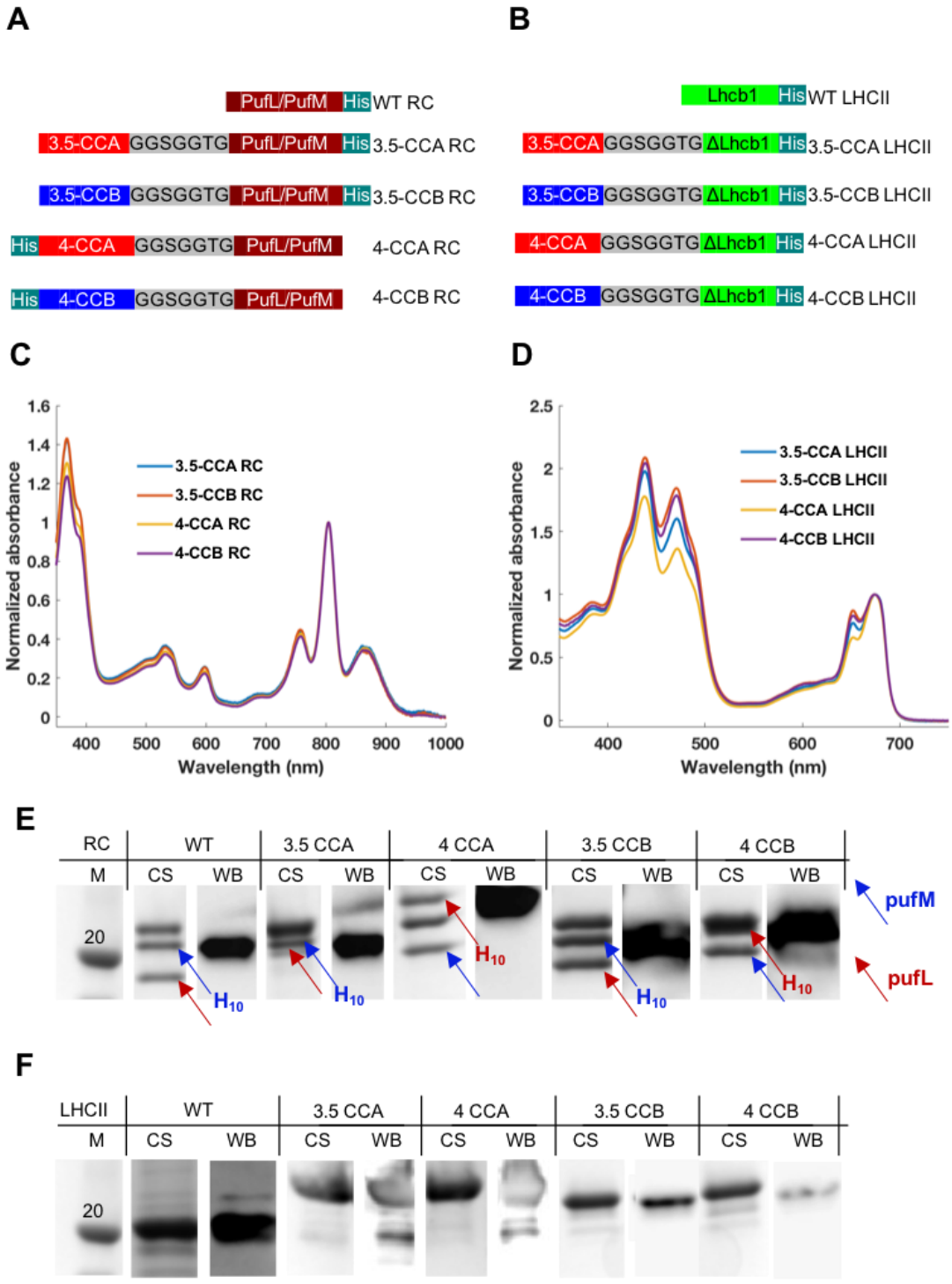


Figure 6.5. Construct design and quality of heterodimeric coiled-coil RC mutants. (A) Coiled-coil RC mutant constructs. (B) Coiled-coil LHCII mutant constructs. The first 11 amino acids of LHCII were replaced by the coiled coil motif and a 7 amino acid flexible linker. (C) Absorption spectra of coiled-coil RC mutants normalized at 804 nm (Q_y of $B_{A/B}$). (D) Absorption spectra of coiled-coil LHCIIs normalized at 674 nm (Q_y of Chl *a*). (E) SDS-PAGE/WB of purified CC-RC mutants with the positions of PufM, PufL and the His-tag indicated by coloured arrows. (F) SDS-PAGE/WB of purified CC-LHCII mutants. CS – Coomassie stain; WB – Western blot with HRP-conjugated anti-polyHis antibody. Uncropped gel images are shown in **Fig. S6.1**-**S6.3.**

6.2.2.2. BN-PAGE screening of coiled-coil RC and LHCII pairs

BN-PAGE was chosen for quick screening for hetero-dimerization between a RC and a LHCII mediated by all eight possible E/K (CCA/CCB) coiled-coil pairs. Because of the high purities of the prepared proteins, most bands that appeared on gels were attributable to either RC or LHCII, and this was subsequently confirmed by in-gel absorption scanning.

The initial screen of the eight possible acid/base coiled-coil combinations is shown in **Fig. 6.6A**. Single protein samples for six of the eight RC or LHCII components are also shown. Bands at the bottom of the gel corresponded to monomeric LHCII and RCs, whilst a band at ~300 kDa corresponded to a RC homodimer, and was seen to varying extents in most RC-only samples and RC/LHCII mixtures. Just below this was a band, particularly prominent in lane 8, that corresponded to a RC+LHCII heterodimer.

A second set of data for six of the eight possible pairs is shown in **Fig. 6.6B**; the loading for this gel was lower to better resolve components. In all cases the predominant bands were an LHCII monomer band at ~60-70 kDa and a RC monomer band at ~160 kDa. Evidence for heterodimerization was seen for two combinations, 4E RC/4K LHCII (**Fig. 6.6B**, Lane 1) and the reverse 4K LHCII/4E RC (Lane 6), with a double band at ~230-240 kDa. Unlike most other pairs, these 4E/4K combinations did not have a band at ~300 kDa attributable to a RC homo-oligomer. These homo-oligomers were particularly noticeable when the RC had a 3.5K or 4K modification (lanes 3,4,5), and were seen in controls without LHCII (lane 7), but were not seen when the 4K RC was mixed with the 4E LHCII (lane 6). In pairs involving a 4K or 4E modified RC there was also a band of varying intensity at very high molecular weight indicative of large aggregates, the position of which changed upon mixing with an LHCII (**Fig. 6.6B**, compare Lanes 5, 6 and 7). The pattern of bands was similar under conditions of 0 mM and 200 mM

NaCl (**Fig. 6.6B**, Lanes 1-7 compared with 8-14), indicating no significant dependence of ionic strength on aggregation state.

The conclusion drawn from this screen was that a low yield of a stable RC≈LHCII (≈ represents a non-covalent linkage) conjugate could only be formed with the 4K/4E pairs, which presumably could be attributed to an insufficient binding force being provided by coiled-coil modifications other than 4K/4E combinations. However, it remains elusive as to why such high affinity was required for the conjugation of RC and LHCII. One issue may have been competition between the interaction between opposing charges required for formation of the designed coiled-coil RC≈LHCII heterodimer and non-specific charge-charge interactions between RCs. It was noticeable that homodimerization seemed to mainly affect RCs modified with a basic coiled coil, which could indicate an interaction between this coiled coil and the predominantly anionic surface of a second RC that was stronger than any basic/acidic coiled-coil interaction except when the 4K coiled coil on the RC was partnered with the 4E coiled coil on LHCII. The possible origin of the larger molecular weight aggregates was less clear, as these were seen both in samples where RC homodimers were apparent and samples where a RC≈LHCII homodimer was formed. It was noticeable that the larger aggregates tended to be seen in samples where the coiled-coil on the RC was a 4-heptad repeat, including in a sample where LHCII was absent, and so this aggregate is mainly likely to represent multiple RCs interacting a non-specific way.

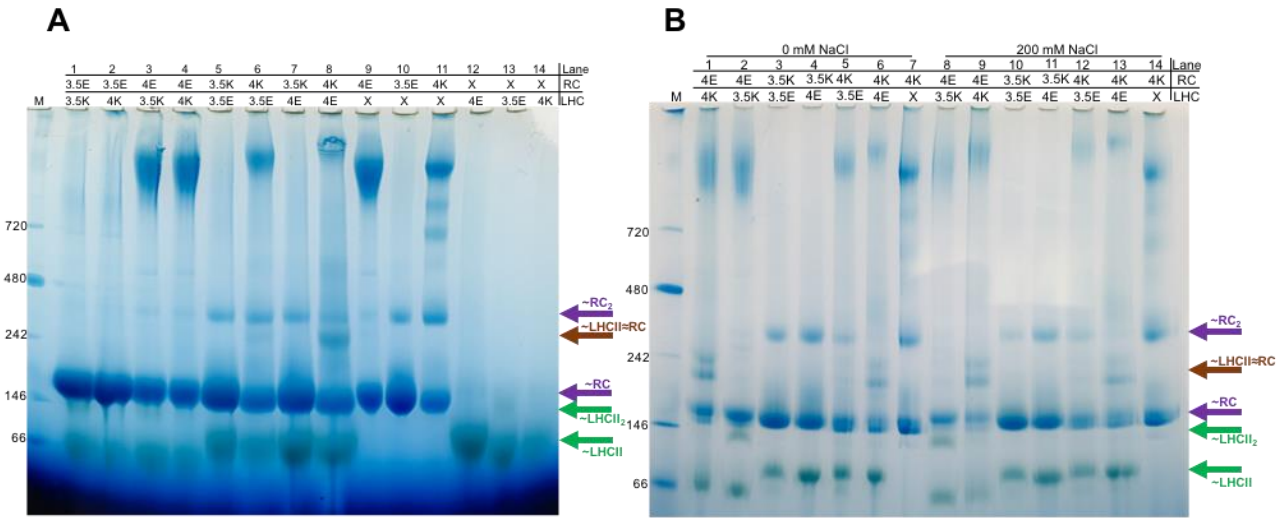


Figure 6.6. BN-PAGE of eight combinations of E/K coiled-coil RCs and LHCIIs to screen for heterodimers. (A) Screen for pairs that form heterodimers. (B) Examination of the effect of ionic strength on coiled coil mediated interactions. Purple arrows – bands with RCs only;

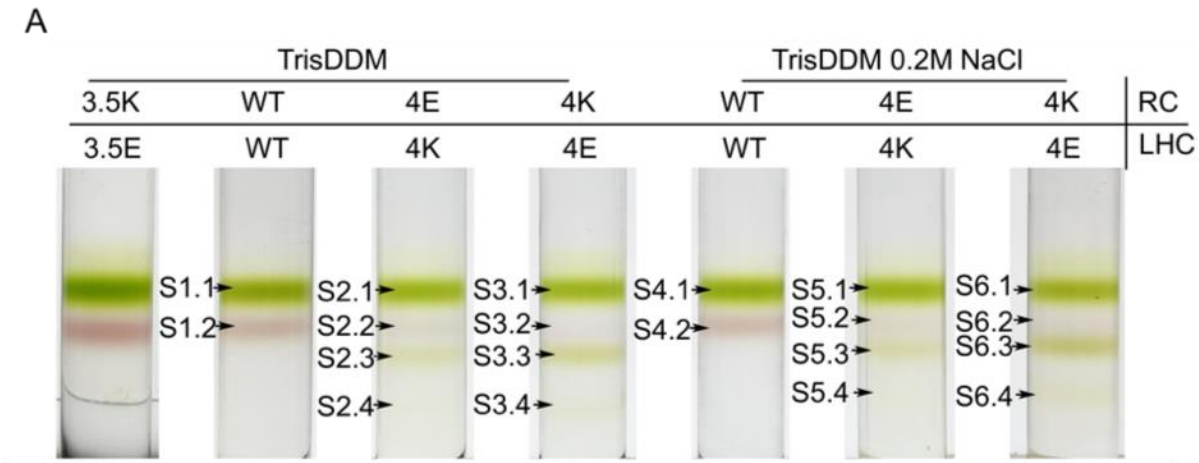
green arrows – bands with LHCII only; brown arrows – bands with LHCII and RC. In sample lane header, the middle row indicates the RC coiled-coil mutant and the lower row the LHCII coiled-coil mutant. X means the absence of RC or LHCII.

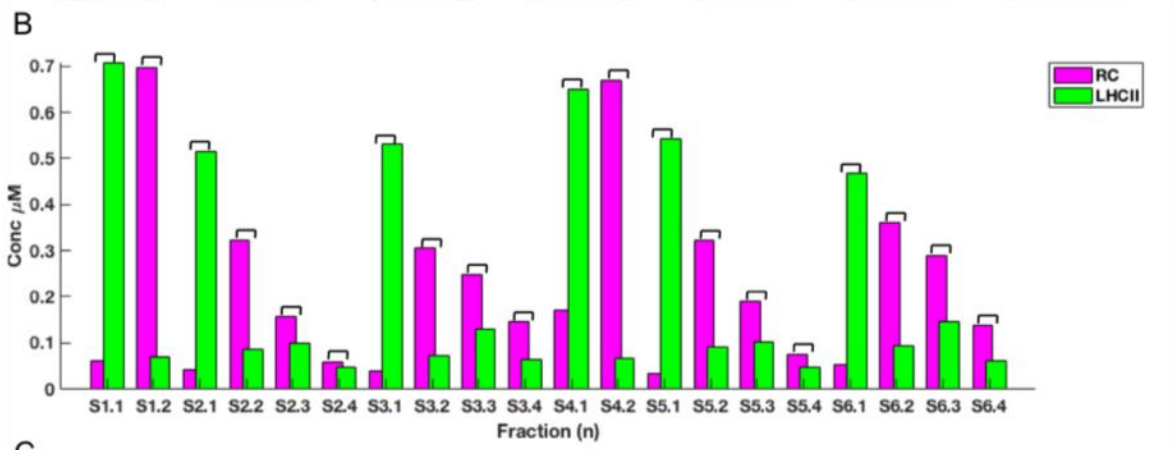
6.2.2.3. Characterization of the stoichiometry of 4K/E RC and LHCII oligomers

Sucrose gradient ultracentrifugation of 4K/E RCs and 4E/K LHCIIs confirmed the presence of RC≈LHCII heterodimers. In comparison to WT/WT and 3.5K RC/3.5E LHCII mixtures, additional bands appeared below the RC monomer for 4K RC/4E LHCII and 4E RC/4K LHCII combinations (**Fig. 6.7A**, bands S2.3 and S3.3 at 0 M NaCl and S5.3 and S6.3 at 0.2 M NaCl) and there was an obvious depletion of the red RC monomer band positioned below the green LHCII monomer band. Absorbance spectra of all the labelled bands in **Fig. 6.7A** showed both RC and LHCII signals after spectral deconvolution of absorption spectra, and computed RC and LHCII concentrations are shown in **Fig. 6.7B**. This analysis consistently gave a >1 ratio of RC:LHCII for the "heterodimer band" (S2.3, S3.3, S5.3 S6.3), varying between 1.6 and 2, which indicated that the compositions of these bands were more complicated than dimer of RC and LHCII (**Fig. 6.7C**). In addition, the presence of faint additional bands at even lower positions in the sucrose gradients (**Fig. 6.7A** bands S2.4, S3.4, S5.4 and S6.4) suggested the RC/LHCII conjugates were heterogeneous.

Two possible reasons for the >1 RC:LHCII ratios in the higher molecular weight bands were considered. One possibility is that the complexes producing these bands were a mixture of closely migrating RC homodimers and RC≈LHCII heterodimers. The second is that some RC/LHC conjugates formed with a greater 1:1 stoichiometry. The double band pattern observed in BN-PAGE (**Fig. 6.7B**, Lane 1, 6, 9, 13) suggested this second possibility might be the better explanation.

Counterintuitively, more intense RC/LHCII oligomer bands were found in 200 mM NaCl/TrisDDM buffer than in just TrisDDM buffer.It was expected that increasing the salt concentration would reduce the specificity between the E/K coiled coils because of a charge shielding effect, but the result with the 4E/4K pairs contradicts this.





C

Band	RC/LHCII ratio	Band	RC/LHCII ratio
S1.1	0.08	S4.1	0.26
S1.2	10.20	S4.2	10.17
S2.1	0.08	S5.1	0.06
S2.2	3.74	S5.2	3.53
S2.3	1.60	S5.3	1.84
S2.4	1.26	S5.4	1.61
S3.1	0.07	S6.1	0.11
S3.2	4.32	S6.2	3.86
S3.3	1.90	S6.3	1.99
S3.4	2.33	S6.4	2.26

Figure 6.7. Sucrose density gradient analysis of 4K/4E coiled-coil RC/LHCII pairs. A 400µL aliquot of each protein (2.5 µM) was used with the exception that 400 µL of 5 µM of 3.5KCC RC and 3.5ECC LHCII was used. (A) images of sucrose gradient. Top row: buffer condition, TrisDDM (20 mM Tris (pH 8.0), 0.04% DDM). Middle row: identities of RC. Bottom row: identities of LHCII. The labels and arrows indicate the 1 mL fraction corresponding to the results in B,C. (B) RC and LHCII concentrations obtained by deconvolution of absorbance

spectra of the indicated fractions/bands. (C) Molar ratios of RC to LHCII in the indicated fractions/bands.

Analytical gel-filtration chromatography was also used to try to gain a better understanding of the 4E/4K coiled coil RC/LHCII oligomers. It was expected that by using an analytical column a better separation of different oligomers could be obtained and the method could be scaled up to purify potential RC/LHCII coiled-coil heterodimers. Collected fractions were assessed by absorbance and emission spectroscopy to gain information on RC:LHCII ratios and the integrity of the LHCIIs. **Fig. 6.8A,B** shows elution profiles for the RC with peaks corresponding to monomers (right) and oligomers (left), whilst **Fig. 6.8C,D** shows equivalent data for LHCII. In each panel the elution profile of a RC/LHCII mix is compared with that of the particular RC or LHCII alone.

The most obvious result from **Fig. 6.8A-D** is that 4K or 4E RCs showed a strong tendency to self-associate whereas the 4K or 4E LHCIIs did not, or at least did so to a much lower extent. A substantial degree of homo-interaction was observed for the 4K or 4E RCs running alone (**Fig. 6.8A,B**) while a small amount of self-aggregation was seen for the 4K LHCII with a double-peak pattern (**Fig. 6.8D**). The elution profiles were changed only subtly when 4K LHCIIs were combined with 4E RCs (**Fig. 6.8B,D**), most obvious change being a small drop in the L:HCII monomer peak and an increase in the higher molecular weight peak (**Fig. 6.8D**). However, more pronounced changes were seen when 4K RCs were combined with 4K LHCIIs, with a shift of the RC oligomer elution peak to somewhat smaller molecular weights (**Fig. 6.8A**) and the appearance of a higher molecular weight elution peak for LHCII accompanied by a drop in the LHCII monomer elution peak (**Fig. 6.8A**). The latter changes became more pronounced when the stoichiometry of 4K RC to 4E LHCII was increased to 2:1.

The general conclusion from this analysis was that the 4K RC/4E LHCII combination worked better than the 4E RC/4K LHCII combination (**Fig. 6.8CD**), which was consistent with the results from sucrose gradients where the “heterodimer band” was more pronounced (**Fig. 6.7A**). In particular the LHCII elution peak shift was clearer for the 4E LHCII than 4K LHCII and the shift showed a positive correlation with the supplied 4K RC (**Fig. 6.8C**).

Due to the efficient pigment network in LHCII, its emission spectrum is independent of excitation wavelength. Therefore, a deviation in emission profile with different excitation wavelengths can be used to evaluate whether the structure of LHCII is distorted by

environmental factors such as associating with a RC. The variance score of normalized emission output was used to screen for potential alterations of LHCII pigment coordination upon coiled-coil mediated conjugation with RCs. This evaluation showed that the emission for LHCIIs conjugated with RCs was not significantly different from LHCII-only controls (**Fig. 6.8E**) with most emission abnormality being attributable to an additional blue-shifted peak under 475 nm excitation. At this wavelength *Chl* b was preferably excited and this emergence of fluorescence on the left shoulder of major emission peak indicated uncoupling of *Chl* b from the total pigment-network[72].

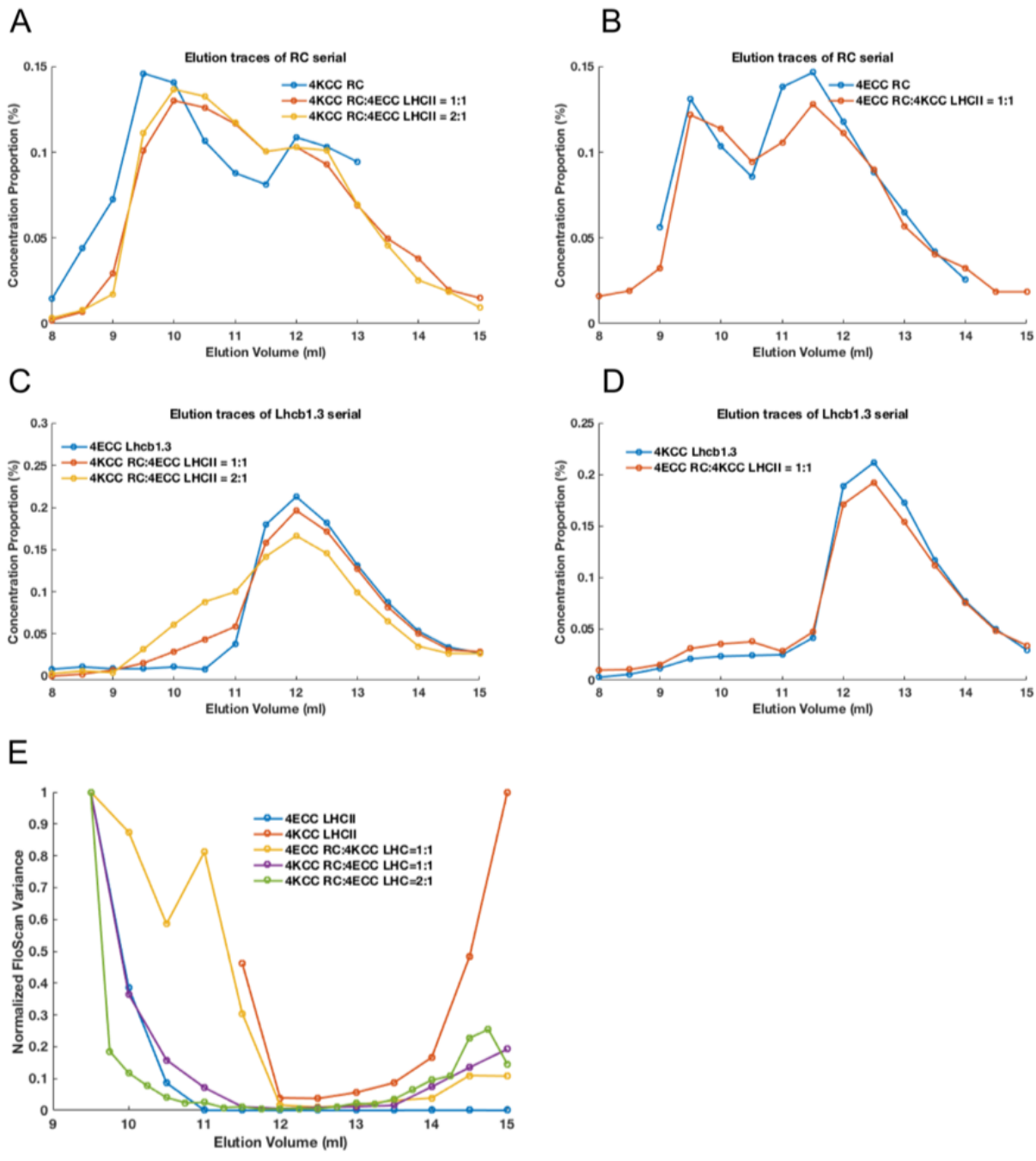


Figure 6.8. Analytical gel-filtration analysis of 4E/4K RC/LHCII pairs. (A) and (B) RC mutant elution profiles with and without LHCII. The concentration of every fraction is given as the proportion of the total concentration. (C) and (D) LHCII mutant elution profiles with and without RCs. (E) Plot of normalized fluorescence variances calculated by integrating the variances of emission scanning with the all possible combinations of emission scans under excitation at 440 nm, 475 nm and 500 nm. To compare tendencies between samples, the values were normalized to the corresponding variance maxima. The higher the score, the more unfolded LHCIIs.

6.2.2.4. Characterization the stoichiometry of 3.5 K-CC RC and 3.5E-CC LHCII aggregates

In previous published work from our laboratory, it was shown to be possible to associate two RCs *in vitro* by modification with a 3.5 heptad repeat CCA/CCB coiled coil. However, in trying to conjugate a RC with a LHCII the combination of these 3.5K and 3.5E coiled coils was not effective, with only 4E/4K combination giving the desired product based on the BN-PAGE screen. As mentioned above it was confirmed using sucrose gradient ultracentrifugation that unlike the 4K/4E or 4E/4K RC/LHCII combinations, the 3.5K/E combination did not produce heterodimers between the RC and LHCII (**Fig. 6.7A**). Unlike the published experiments on 3.5K/3.5E RC heterodimers, which were carried out with LDAO as the supporting detergent, these latter experiments on RC/LHCII heterodimerisation were carried out in the presence of DDM as LDAO is strongly destabilising for LHCII. This raises the possibility that coiled-coil heterodimerisation of these membrane proteins is sensitive to the type of detergent used.

An issue with the use of DDM, which is the most widely used of a relatively small group of detergents in which LHCII is stable, is that it forms rather large detergent micelles. To explore alternatives, the possibility of using amphipols was considered. Amphipols are group of amphiphilic polymers that can associate tightly with the hydrophobic region of membrane protein and displays its hydrophilic group to the surrounding solvent. This makes membrane proteins water soluble in a manner similar to detergents[246]. Amphipol A835 was used because it should form smaller complexes with RCs and LHCIIs compared to those formed by DDM. It should also not show dynamics as high as those shown by micelles comprising small detergent molecules. Either might help a stable coiled-coil mediated interaction to be detected.

The result, as shown in **Fig 6.9A** (and gradient absorbance scans in **Fig S6.5)**, indicated that the formation of a larger conjugate was possible on mixing 3.5K RCs and 3.5E LHCIIs

with an amphipol. However, the ladder pattern suggested various combinations of RCs and LHCIIs were formed.

When the same experiment was carried out with the relatively small detergent octyl glucoside (OG), almost all LHCIIs and RCs migrated nearly to the bottom of the sucrose density gradient, indicating the formation of very large aggregates that were stable in solution (**Fig 6.9B,** with gradient absorbance scans in **Fig. S6.6,** upper three panels). The formation of aggregates was rather independent of the concentration of RCs and LHCIIs, which means that everything was interacting with everything. When mixing the two proteins in the presence of 200 mM NaCl, the aggregate size reduced in general across all three samples, these being 3.5K RCs or 3.5E LHCIIs and a mixture of the two (**Figure 6.9B,** with gradient absorbance scans in **Fig. S6.6,** lower three panels). The effect was similar to that of salt concentration on the RC AR mutants (see above).

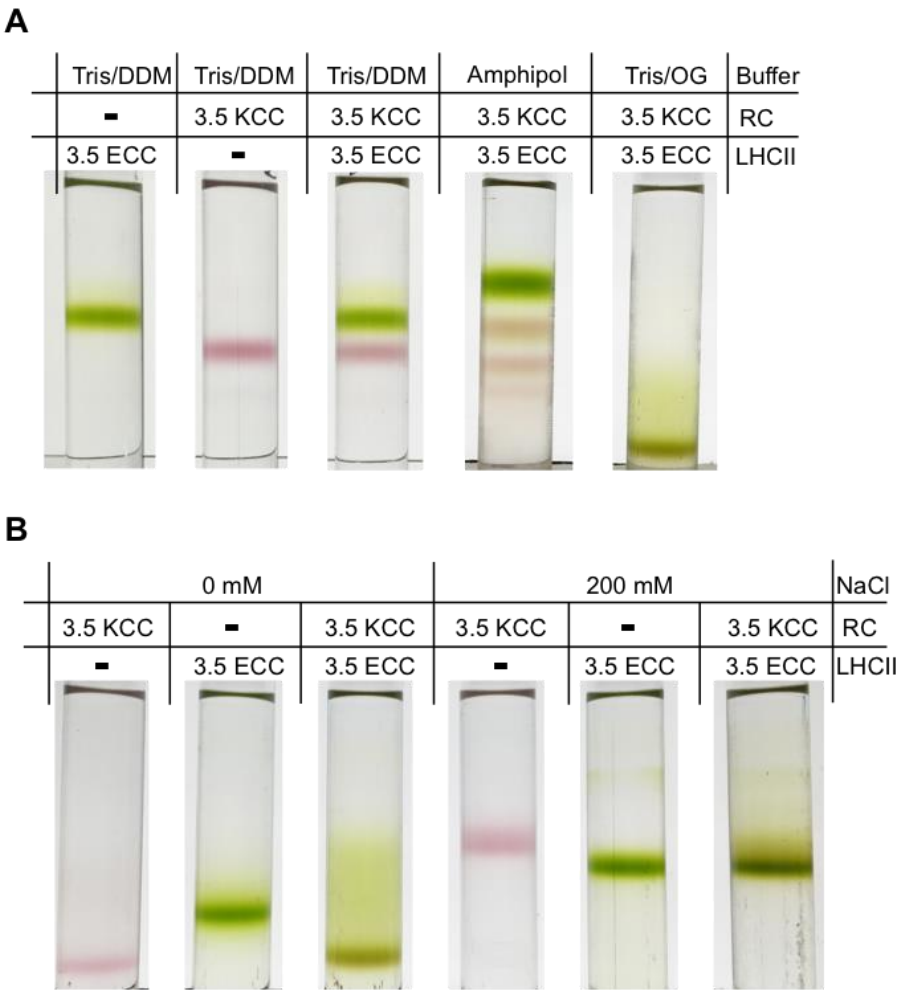


Figure 6.9. Effects of buffer/detergent conditions on heterodimer formation. (A) Effects of amphipol A835 and OG. (B) Effect of ionic strength. Buffer was 50 mM Tricine (pH 7.0)/0.88% OG.

6.2.2.5. Conclusions regarding the use of coiled coils to assemble photoprotein oligomers

The conclusion arrived at from investigation of the use of pairs of acidic and basic helices to drive photoprotein dimerization through the formation of heterodimeric coiled coils was that this approach did not reliably product an acceptable yield of designed product. Although heterodimers could be assembled using pairs of four-heptad repeat acidic and basic α-helices capable of forming a heterodimeric coiled coil, their yield was low, and there was significant formation of alternative oligomeric structures plausibly including RC homo-dimers and larger mass aggregates. As synthesised peptides these designed α-helices produce predictable products with high specificity and binding affinity. However, when attached to large integral membrane pigment-proteins they have limited effectiveness in driving heterodimer formation in solution. Some modulation of the yield of higher molecular weight products was possible through variation of ionic strength and detergent conditions, but they remained a minor product of mixing compatible partner proteins. It was concluded that coiled coils in themselves were not a practical solution to the problem of how to reliably assemble programmed heterodimers with a high yield, and this fed into the concept of the SpyCoil interface described below.

6.2.3. Evaluation of heterodimerization of RC/LHCII by the Cfa intein

Inteins are domains in a family of naturally occurring self-splicing proteins[166]. These proteins have three domains, two exteins and one intein, and the intein domain can be removed from protein sequence using two key cysteine residues in a three-step reaction. The result is that a new peptide bond is formed between the flanking extein domains. Protein engineering has created novel protein linking systems by splitting natural intein domains in two and attaching the two parts to heterogeneous target proteins or short peptides. On mixing the modified components in solution the complete intein domain folds and catalyses the splicing reaction, covalently joining the two proteins or peptides together. The advantage of the split intein approach is that the additional motifs on modified proteins are eradicated from the product fusion protein with the exception of a short tripeptide “scar”.

6.2.3.1. Modification of RC and LHCII sequences with the Cfa split intein domains

To explore this approach to membrane photoprotein fusion, a newly invented highly efficient intein domain (Cfa) was used (see above). The constructs used are shown in **Fig. 6.10** and in **Table S6.1**. The N-terminal domain of the split intein (denoted as CfaN) was fused to the C-terminus of PufM of the RC and the C-terminal polypeptide of the split intein (CfaC) was inserted at the N-terminus of LHCII. As can be seen from **Fig. 6.10** a His tag was placed adjacent to the CfaN or CfaC domain to ensure it was retained in the final purified RC or LHCII, the CfaN and CfaC domains included a C or CFN sequence required for intein splicing, and the domains were attached to the target protein by a 7 amino acid flexible linker. The scar left by this split intein system has the sequence CFN, and in the final fusion protein will be flanked by the two flexible linkers giving a 17 amino acid linking sequence between the RC and LHCII. CfaC- and CfaN-modified eGFP constructs were also designed as control proteins (**Fig. 6.10)**.

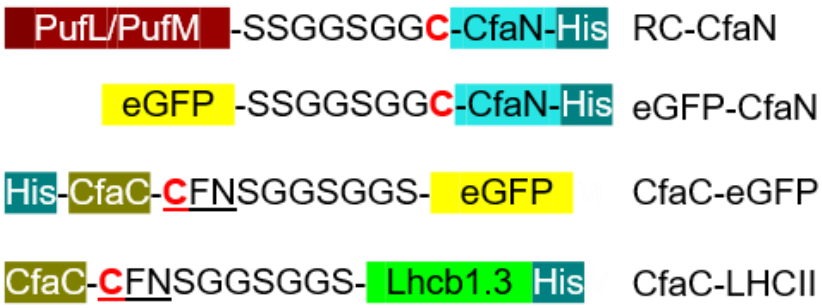


Figure 6.10. Cfa mutant constructs. In the RC-CfaN construct the CfaN domain was attached to the C-terminus of PufM via a linker and before a His-tag. In the eGFP-CfaN construct the CfaN domain was attached to the C-terminus of eGFP via a linker and before a His-tag. In the CfaC-eGFP construct the CfaC domain was added to the N-terminus of eGFP via a linker preceded by a His-tag. In the CfaC-LHCII construct the CfaC domain was attached to the N-terminus of Lhcb1.3 via a linker and a His-tag was present on the C-terminus of Lhcb1.3. More information can be found in **Table S6.1**. In all constructs the key cysteine used for the intein splicing reaction is highlighted in red.

6.2.3.2. Heterodimerization of RC and LHCII with a split intein

RC/LHCII heterodimerization was tested by mixing purified RC-CfaN and CfaC-LHCII proteins in solution, and test reactions were also run with a mixture of RC-CfaN and CfaC-eGFP. The overall outcome of these tests was a frustratingly low yield of the desired product

and high yields of “abortion” products that are known to be a possible issue when applying the split intein approach[167]. When working efficiently, the reaction scheme should follow the vertical route in **Fig. 6.11A**, with transesterification connecting protein P1 to protein P2 via the CFN scar. However, two side reactions are also possible that either cleave the CfaC domain from protein P2 (Abortion 1) or cleave the CfaN domain from protein P1 (Abortion 2).

Experiments on use of the Cfa split intein with the RC and LHCII showed a substantial amount of side reaction activity that strongly limited the yield of corrected covalently linked product. A reaction between RC-CfaN and an excess of CfaC-eGFP gave limited production of the target RC-eGFP heterodimer (labelled as “Product” in **Fig. 6.11B**, see original gel in **Fig. S6.7**). Analysis of band intensities revealed that for every successful RC-eGFP dimer formed there was a five-fold greater yield of Abortion 1 product at room temperature or three-fold greater yield at 4 ^{o}C as the result of processing of the CfaC domain off the N-terminus of eGFP (**Fig. 6.11B**, 3rd and 4th lane). A similar result was seen in the RC-CfaN/CfaC-LHCII reaction with large amount of Abortion 1 product (**Fig. 6.11C**, see whole gel in **Fig. S6.8**) and only a very low yield of a RC-LHCII fusion protein.

Therefore, it appeared to be that intein reaction was sensitive to the size of the attachments to one or both of the CfaC and CfaN domains. This might be attributed to the weakening of the association between the spliced intein domains when attached to high molecular weight proteins. When the corrected alignment of the split CfaN/CfaC intein was disturbed, the unstable thioester bond could be hydrolysed and give abortive products with losses of CfaN or CfaC depending on where the breakdown happened (**Fig. 6.11A**).

The products of a scaled-up linkage reaction between RC-CfaN and CfaC-LHCII were analysed by sucrose gradient ultracentrifigation (**Fig. 6.11D**). This produced a yield of RC-LHCII dimer of less than < 10%, with evidence of loss of the CfaC domain from LHCII when bands from the gradient were analysed (**Fig. 6.11D**, see whole gel in **Fig. S6.9**). Attempts were made to reduce interference from side reactions by varying the reaction conditions such as buffer, temperature or the starting concentrations and the ratio of CfaN/CfaC proteins, but little improvement was observed.

6.2.3.3. Conclusions regarding the use of split inteins to assemble photoprotein oligomers

Although, in principle, the split intein approach is an attractive way to covalently lock two integral membrane proteins together, in practice this approach was not practical given strong

interference from side reactions. It is possible that a different type of split intein might give better results, as might modified construct designs with longer linkers, but there was insufficient time in the project to explore such options.

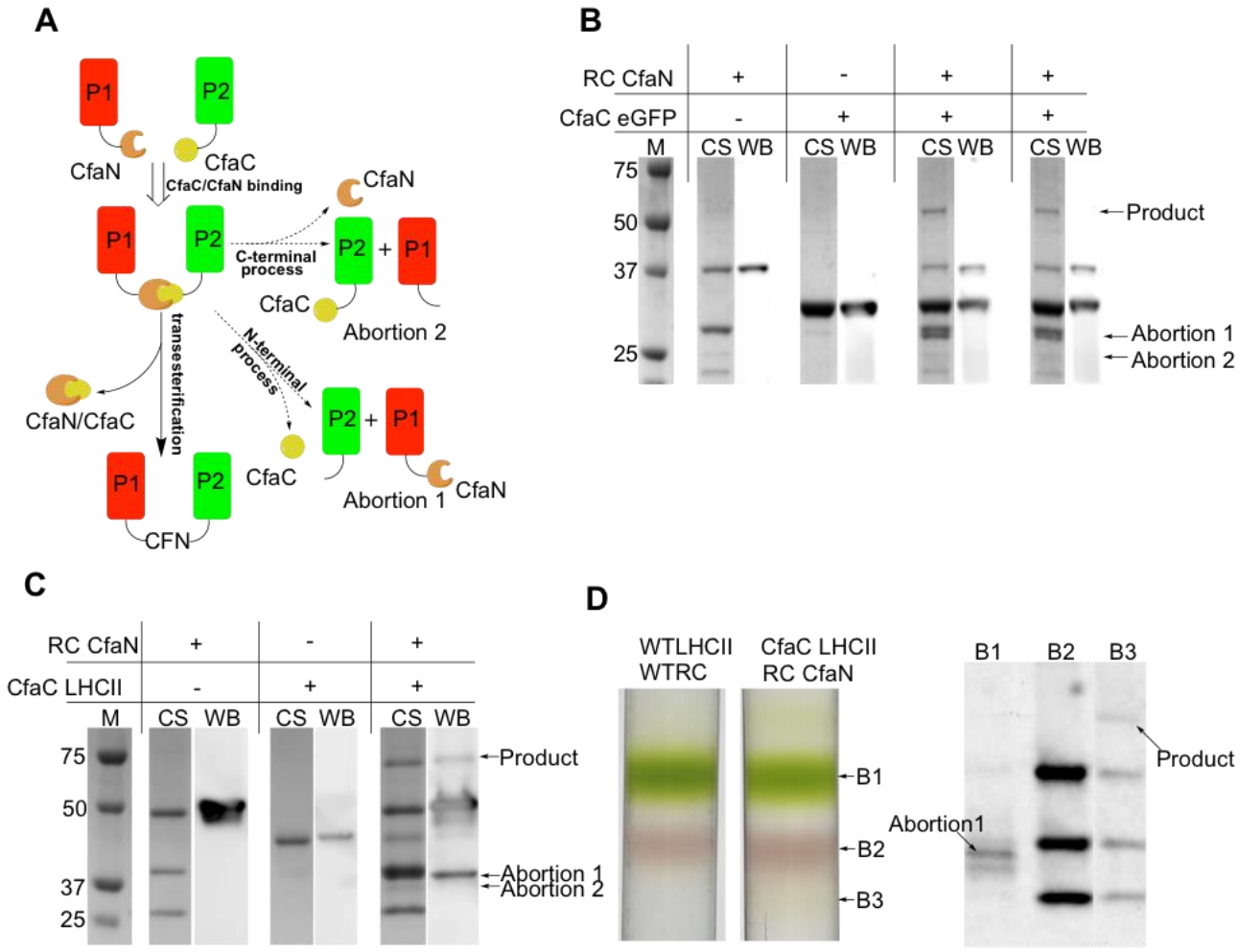


Figure 6.11. Construction of a RC/LHCII heterodimer using intein Cfa. (A) Scheme of then intein splicing reaction between two proteins modified with split Cfa (CfaN and CfaC). Abortion1 and Abortion 2 can happen if the transesterification is alleviated via hydrolysis of either the N-terminal or C-terminal thioester bond. (B) Reaction between RC-CfaN and CfaC-eGFP with the successful and abortive products labelled with arrows. The reaction was carried out in buffer with 5 mM TCEP overnight. The 3rd lane shows the products of a reaction at room temperature and 4th lane the products of a reaction at 4 ^{o}C. (C) Reaction between RC-CfaN and CfaC-LHCII l with the desired and abortive products indicated. (D) Separation by sucrose density gradient ultracentrifugation of the products of a scaled-up split intein reaction, and SDS-PAGE/Sypro Ruby staining of fractions corresponding to bands B1, B2 and B3.

6.2.4. Design of SpyCoil and its utilization for mega-protein assembly

In addition to testing three published methodologies for covalently or non-covalently associating two normally foreign pigment-proteins, namely SpyTag/SpyCatcher, coiled coils and split inteins, a new hybrid approach was developed and tested. This methodology, dubbed "SpyCoil" was designed as a combination of the P3/P4 heterodimeric coiled coil, one of three designed orthogonal pairs[148] and the SpyTag/KTag/SpyLigase system[149], the idea being to bring the advantages of the two linking strategies together.

6.2.4.1. Rationale behind the design of SpyCoil

The design of SpyCoil stemmed from the idea of improving the activity of the SpyTag/KTag/SpyLigase system with an associated heterodimeric coiled coil to bring the SpyTag and KTag elements into close contact in solution. As outlined above, the KTag component was derived from the SpyCatcher domain by further excising the β-strand containing the reactive lysine to create a KTag peptide and a residual SpyLigase domain. It was shown that on mixing the SpyTag, KTag and SpyLigase components an isopeptide bond was formed between SpyTag and KTag but only in a specific buffer containing 1 M TMAO[149]. This is because a deterioration in binding affinity occurred following over-splitting of the natural adhesin domain into too many pieces. Hence one rationale behind the design of SpyCoil was to bring affinity back by providing an additional binding domain, while the *de novo* designed coiled coils also introduce specificity to the system. Exploiting coiled-coil affinity/specificity also opened up the possibility of creating many orthogonal SpyCoils for constructing as many complicated assemblies as the design space allowed.

An important consideration was that the position of the coiled-coil bundle should be close to the SpyTag and KTag sequences but not so close that it blocks the formation of the SpyTag/KTag/SpyLigase complex. Therefore, the design process focused mainly on the length of flexible linkers to join the SpyTag and KTag peptides to the two halves of the heterodimeric coiled coil. The SpyTag sequence was combined *in silico* with that of the P3 (basic) coiled-coil peptide whilst the KTag sequence was combined with that of the P4 (acidic) coiled-coil peptide.

6.2.4.2. *In silico* design of SpyCoil

Starting with SpyTag, one glycine was appended to the C-terminus of SpyTag, followed by the P3 sequence, and free sampling of all possible dihedral angles (φ and ψ) around the SpyTag/glycine peptide bond was carried out (**Fig. 6.12A**; SpyTag-P3 shown in blue). The

initial structure comprised the SpyTag-P3 fusion plus separate KTag and P4 (**Fig. 6.12A**; KTag and P4 shown in blue). All possible configurations were scored with BUDE[160] (Bristol University Docking Engine) and it was found that about 94% had a score of < -50, indicating a good degree of freedom between the two domains (**Figure 6.12B**). Including SpyLigase in the query (**Fig. 6.12A**; green), the BUDE score showed that 92% of the permitted states were also allowed upon formation of a complex with SpyLigase in terms of probability deduced from a Boltzmann distribution considering that BUDE score was a direct representation of total system free energy (**Fig 6.12B**).

Since a good freedom of states was achieved by introducing one glycine between SpyTag and P3 sequence, the next step was to join KTag to the P4 sequence. State 72 was chosen from the first stage of the analysis because it presented the longest possible distance between the C-terminus of KTag and N-terminus of P4 among all of the queried conformations (**Fig. 6.12A,** state 72). The idea was that if a KTag-P4 linker allowed this most extended configuration, then the other less extended conformations could be also adopted, ensuring little reorientation energy input was required upon complexing with SpyLigase. Meanwhile, the linker should not be too long to attenuate the benefit the coiled coil brought to the system in terms of bringing SpyTag and KTag close together. Finally, possible linkers were designed in Modeller using the loop refine tool. Five refined loops were generated ranging from 6 amino acids (AAs) to 10 AAs and the average discrete optimized protein energy (DOPE) scores were estimated at the connection points between each flexible linker and the coiled-coil sequences (**Fig. 6.12C**). It was found that the 9 AAs linker was the best option as its DOPE score was low and the score seemed to plateau with increasing linker length beyond this point. Therefore a 9 AAs linker was taken for later stages of molecular dynamics (MD) verification and experimental testing.

6.2.4.3. Molecular dynamics study of SpyCoil

The designed SpyCoil construct was verified by MD to check whether there were any unnoticed structural obstructions. A 50 ns simulation showed, upon formation of the full “SpyComplex” with SpyLigase, SpyCoil experienced no large conformational change (**Fig. 6.13A**, top row), as both the root-mean-square deviation (RMSD - **Fig. 6.13B**) and radius of gyration (Rg - **Fig. 6.13C**) remained nearly constant across the simulation time. As illustrated in **Fig. 6.13A** (second row), in the absence of SpyLigase, SpyCoil folded up as indicated by an increased RMSD and reduced Rg (**Fig. 6.13BC**, labelled “TwoTags”). The two termini of the SpyCoil sequences were generally more flexible especially in the absence of SpyLigase,

which indicated by root mean square fluctuation (RMSF), was reduced when forming a complex with SpyLigase, indicating a stable complex formation (**Fig. 6.13DE**). In conclusion, it was concluded from the MD that the design of SpyCoil was physically feasible.

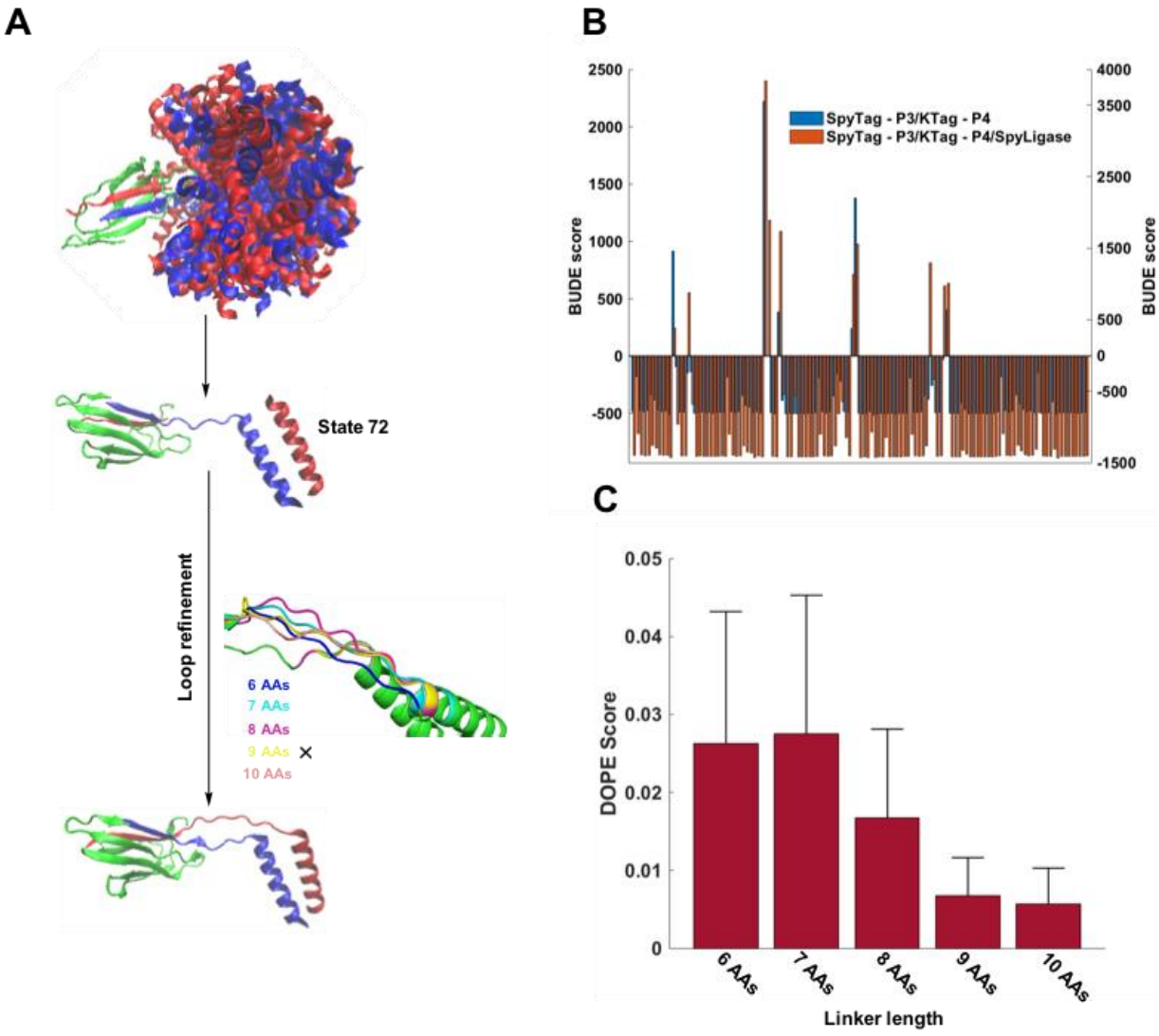


Figure 6.12. *In silico* design of SpyCoil. (A) Scheme of design work flow with the model structure of SpyLigase (green), SpyTag-P3 (Blue) and KTag-P4 (Red). The top image depicts all permitted dihedral angles around the flexible glycine joint between SpyTag and P3. (B) BUDE scores of all configurations of SpyTag-P3 + KTag-P4 with or without the inclusion of SpyLigase. Small negative or positive scores suggested backbone clashes in the queried complex structures. (C) Average DOPE score of the designed loop from 6 to 10 AAs long (n = 5). Only the last two amino acids of the linker and the first two residues of P4 were considered as other parts of the polypeptides were not changed.

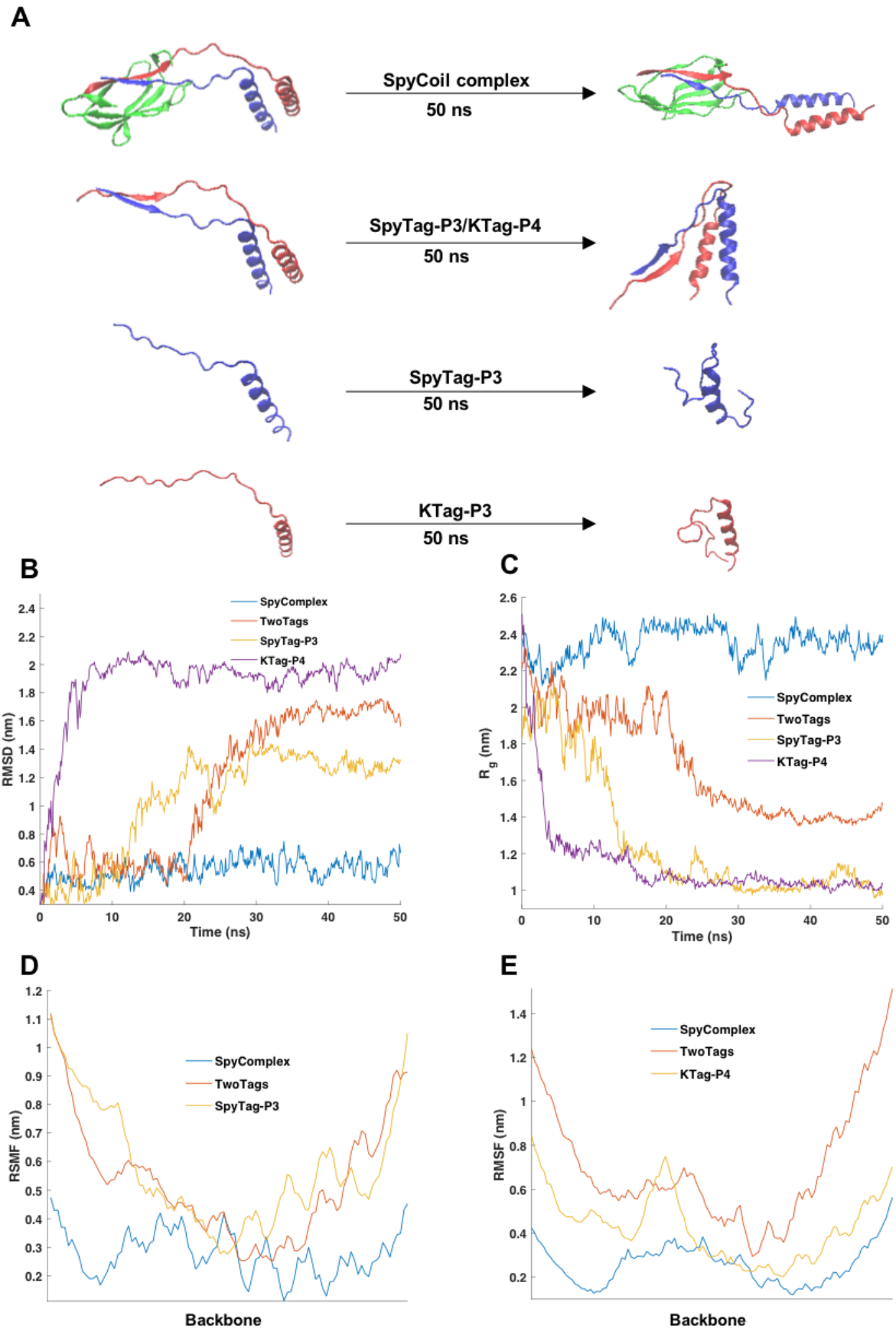

A
SpyCoil complex
50 ns
SpyTag-P3/KTag-P4
50 ns
SpyTag-P3
50 ns
KTag-P3
50 ns
B
SpyComplex
TwoTags
SpyTag-P3
KTag-P4
RMSD (nm)
Time (ns)
C
SpyComplex
TwoTags
SpyTag-P3
KTag-P4
Rg (nm)
Time (ns)
D
SpyComplex
TwoTags
SpyTag-P3
RSMF (nm)
Backbone
E
SpyComplex
TwoTags
KTag-P4
RMSF (nm)
Backbone

Figure 6.13. Molecular dynamics analysis of SpyCoil constructs. (A) Presentation of model structures at the start and end state after a 50 ns simulation. SpyLigase, SpyTagP3 and KTagP4 are shown in green, blue and red respectively. (B) Root-mean-square deviation (RMSD) of whole estimations. As expected, the two linkers were more flexible in the middle of SpyCoil with higher RMSF. (C) Radius of gyration of whole protein backbones versus simulation time. (D) Root mean square fluctuation (RMSF) of SpyTagP3 backbone atoms when aligned with SpyLigase and KTagP4 as a SpyComplex or with KTagP4 as TwoTags or as the lone component of the simulation. (E) The RMSF of KTagP4 backbone atoms when aligned with SpyLigase and SpyTagP3 as a SpyComplex or with SpyTagP3 as TwoTags or as the lone component of the simulation.

6.2.4.4. Evaluation of the functionality of the designed SpyCoil linker

For evaluation, the designed SpyCoil sequences were synthesised as fusions to either terminus of model protein eGFP, as shown in **Fig. 6.14** and **Table S6.1**. By conjugating SpyTag and KTag to the heterodimeric P3/P4 coiled coil, SpyLigase activity in physiological buffer was restored and the reactions could be completed with any combination of SpyCoil eGFP. The formation of fusion protein product was assessed by PAGE (**Fig 6.15A**). The appearance of bands around the 75 KDa marker was due to covalently-associated eGFP dimers being formed in the presence of SpyLigase (the expected molecular weight was 67 KDa). No 75 KDa band could be detected in the presence of a catalytically inactive mutant SpyLigase (labelled EQ)[149], or in test reactions with single proteins. Under conditions of SpyCoil-eGFP:SpyLigase = 5:1 and [SpyCoil eGFP] = 5 μM, the yield of product was uniformly above 90% after incubating at 4^{o}C overnight (**Fig 6.15A**). A point to note was that although the SpyTag-P3 and KTag-P4 sequences were connected with different orientations at the N-terminus and C-terminus of eGFP (**Fig. 6.14**), surprisingly these could react in any combination without any effect on the ligation efficiency (**Fig 6.15A**). Interestingly, the high completion of the reaction with C-terminal SpyCoils signified that one can lock two proteins with the isopeptide bond at a proximal position with SpyTag/KTag closest to the host proteins.

Using purified SpyCoil-eGFP proteins directly after nickel purification with 0.5 M imidazole, dimerization of eGFP mutants could be still obtained (**Fig. 6.15B**, and see original gel in **Fig. S6.10**). The addition of 0.5 M NaCl lowered the reaction yield dramatically whereas it was demonstrated by Zakeri *et al.* that Spy-based isopeptide bond formation was not sensitive to buffer conditions[147]. This indicated that the isopeptide bond formation depended on coiled-coil dimerization as P3 and P4 carried strong counter charges at the ***g*** and ***e*** positions[148].

It was also shown that the SpyCoil linking strategy was also plausible between eGFP and an LHCII modified with KTag-P4 at its C-terminus (**Fig 6.15C**, and see original gel in **Fig S6.11)**. All in all, it seemed SpyCoil was a great solution toward the design and construction of stable mega-assemblies involving many different proteins with very precise control over the assembly order and final configurations as along as it can be stably expressed with the host proteins.

Figure 6.14. Constructs for designed SpyCoil proteins. The constructs were attached to either the N- or C-terminus of eGFP creating four combinations. One additional LHCII-KTagP4 construct was included to demonstrate the feasibility of the SpyCoil strategy with membrane photoproteins. More information can be found in **Table. S6.1**.

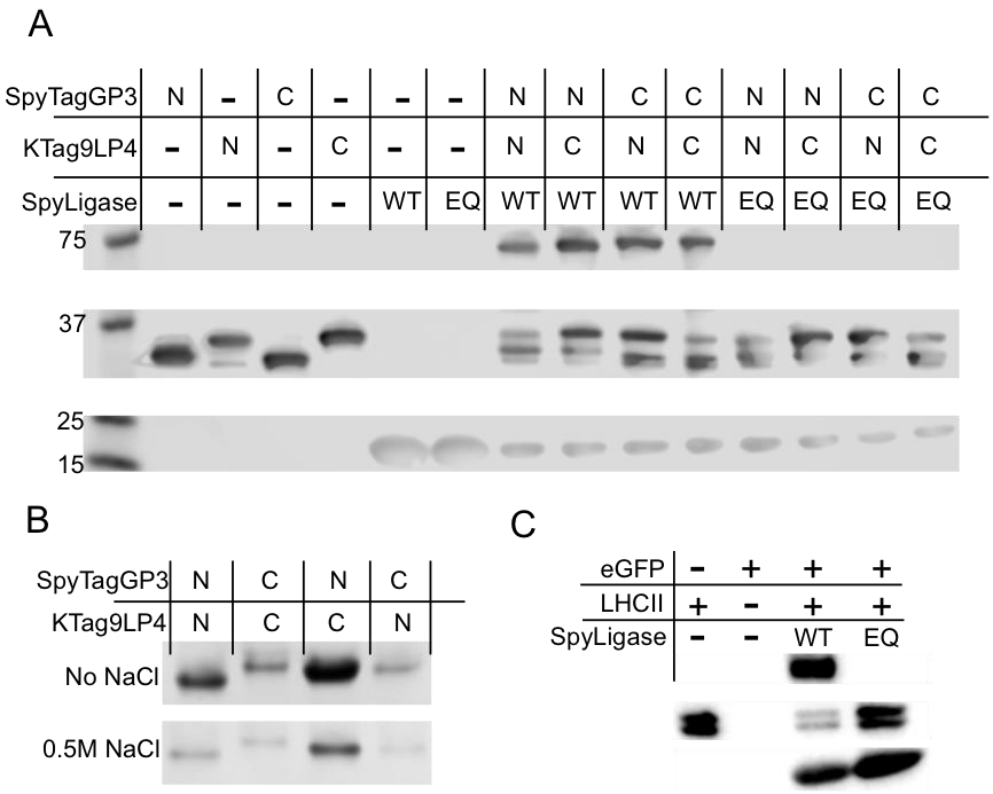


Figure 6.15. SDS-PAGE/WB analysis on eGFP dimerization with SpyCoil. (A) WB results of SpyCoil ligation at all combinations of terminals. SpyCoil reactions were setup with SpyLigase or SpyLigase EQ (inactive mutant). The formation of covalently locked eGFP dimers were seen by the appearance of bands slightly below 75 KDa comparing to the calculated molecular weight of dimerized products are about 66 KDa. "-" signified the component it represented is not included into the reaction mixture. N and C represent the position of the SpyCoils peptide at either N-terminus or C-terminus. WT indicates active SpyLigase and EQ an inactive SpyLigase with a point mutation on key glutamate. The reactions were all carried out in TrisDDM buffer. (B) Test of SpyCoil under different buffer conditions. WT SpyLigase was supplied with 0.5 M imidazole present in the buffer with or without 0.5 M NaCl. (C) WB showing SpyCoil can covalently attach eGFP to LHCII with an active SpyLigase. "+" indicated the presence of represented component in the reaction mixtures.

6.3. Discussion

In this chapter, different linking strategies either coming from the literature or developed in-house were tested for bottom-up construction of photosynthetic membrane assemblies. Although the published post-translational protein association strategies have been reported to be effective for constructing other conjugates, they were not effective in the present system either because of a shortage of precise control, low efficiency, side reactions or incompatibility with host proteins or experimental conditions.

Regarding coiled coils, it was established that use of a 4 heptad-repeat E/K coiled coil pair did produce some RC≈LHCII heterodimers, but it mostly led to self-association of RCs in particular, and LHCII to a lesser extent. In addition, RCs modified with double homodimeric coiled coils could self-assemble into large aggregates, initialization of which could be controlled by buffer conditions. In general, the formation of higher molecular weight complexes was also sensitive to salt concentration and choice of detergent. The principal conclusion was that with the CCA/CCB coiled coils the oligomeric state of protein oligomers was rather hard to control, and that *de novo* designed coiled coils seemed not to be the general solution for building protein assemblies. Many of the published applications of these counter-charged coiled coils do not involve very detailed interpretation of the architectures of the resulting complexes and usually do not report the stoichiometry of the designed complexes. It was recognized here and by other research groups that both CCA and CCB have a tendency to homodimerize, with CCB tending to homodimerize more than its counterpart due to the lysines used to determine the specificity at the ***g***,***e*** position tolerating more homo-packing because of their long side chain[247]. Although currently available coiled coils, such as the CCA/CCB pair, can in principle be used to construct multi-protein complexes, due to undesirable features such as the trade-off between affinity and specificity, restriction in length and potential influence on natural structure, coiled coils may not be the best option for molecular level control of protein assemblies in a modular approach. Hence more effort needs be made on designing bespoke coiled coils for specific tasks and perhaps the rational design of coiled coils should be carried out "*in situ*" to take potential influences from different environments into account.

Split inteins, merited for their high efficiency, short scar and controllable reactions, are one of the most attractive protein linking strategies. Split inteins can link desired protein domains together with very short scar, a feature that is very different from many other approaches that leave large adducts on the final products such as a coiled-coil bundle or the SpyCatcher/SpyTag domain. The key cysteine must be in the reduced state and therefore the

reaction can be controlled through the reducing power of the environment. The intein self-splicing reaction is relatively fast and completes with small adducts. However, attempts to utilise the Cfa split intein in the present work showed that the system could not handle large membrane protein/micelle complexes as efficiently as was originally reported with small peptides. The final product of the conjugation could be precisely identified as a covalently-locked heterodimer of RC and LHCII. However, the yield was no more than 10% with significant amounts of side reaction products, which indicated that the trans-peptidyl reaction was being interfered with by the very large attachments presumably due to a weakening of the stability of the intein complex. Despite these frustrations, split inteins are certainly the cleanest way to construct large protein assemblies and there are many members of this family of proteins. Effort should be made to find better intein candidates and splitting strategies, or protein engineering should be explored to enhance the complex stability to minimize the abortive reactions. The power to control the reaction through redox potential might be interesting for smart control of assemblies such as production of reductants in photosynthesis, which could be adapted as a means to control photochemistry by manipulating photosystem organization.

The final newly-designed peptidyl-linker, SpyCoil, seemed to be the most promising alternative linking strategy to SpyTag/SpyCatcher due to its stability and the easy creation of orthogonal pairs. It combined the merits of the pairwise design of coiled-coil heterodimers and ultra-stable isopeptide locking from the Spy-system. The lack of activity of SpyTag/KTag/SpyLigase under most “standard” buffer conditions may at first be considered as a disadvantage. However, it offered great design space for specificity so that one can create as diverse pairs as possible. As demonstrated here with the P3/P4 coiled coil pair, the activity of SpyTag/KTag/SpyLigase could be restored by providing additional affinity to the two tags. In addition, owing to the well-characterized design principle and parameterization of the coiled coil motif, one can readily create many orthogonal pairs of SpyCoils and be capable of simultaneous self-assembly of proteins, or even other components modified with SpyCoils, into desirable configurations with excellent stability. Due to the isopeptide bond formed between paired SpyCoils, the off-targeting associated with coiled coils could be suppressed to some extent, easing design restrictions as well. Meanwhile, the created multi-protein complexes would be ultra-stable, which should benefit their application in many fields such as biohybrid solar cells, biosensors, therapeutics and vaccination. Another advantage of SpyCoils, as with other isopeptide-based linkers (e.g. SnoopTag/SnoopCatcher), is their compatibility with living systems. Since the SpyCoils can be modified onto a protein of interest and expressed inside cells, they can be deployed *in vivo* and bring the design of complicated cell function, such as photosynthesis, into reality. One example would be the design of plants

or algae with enhanced photosystem productivity or light tolerance by assembling desirable enzymes into clusters or manipulating photosynthetic membrane configuration.

6.4. Conclusions

To conclude, it was demonstrated in this chapter that various linking strategies can achieve conjugation of photoactive proteins. However, the *de novo* designed heterodimeric coiled coils and the split intein systems tested gave only a very low yield of desired product, and neither presented a practical alternative to the SpyTag/SpyCatcher system utilised in Chapter 5. A newly-designed SpyCoil system showed more promise, although a lack of time during the project prevented full testing of its capabilities with more than one integral membrane photoprotein. Encouragingly, augmenting the SpyTag/KTag system with associated coiled coils greatly improved the buffer conditions under which SpyLigase is able to function, and the system showed good flexibility with regard to where the Tag/Coil peptides could be fused to test proteins.

6.5. Supplementary figures

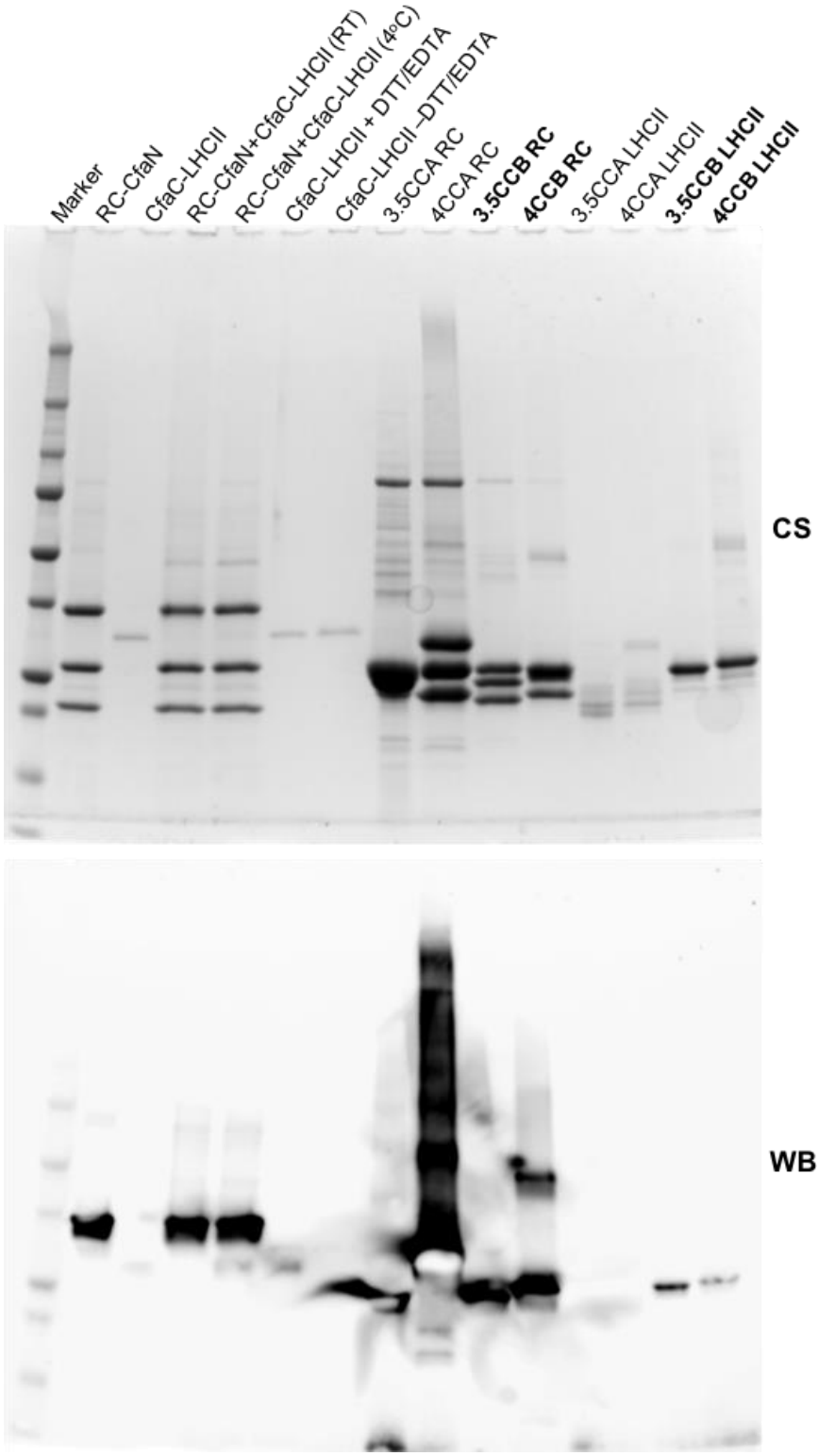


Figure S6.1. Original image of SDS-PAGE and Western blot. Lanes used in **Figure. 6.5EF** are highlighted in bold. CS – Coomassie Stain; WB – Western Blot. The A700 channel was

merged with the WB signal for simultaneously visualizing the protein marker.

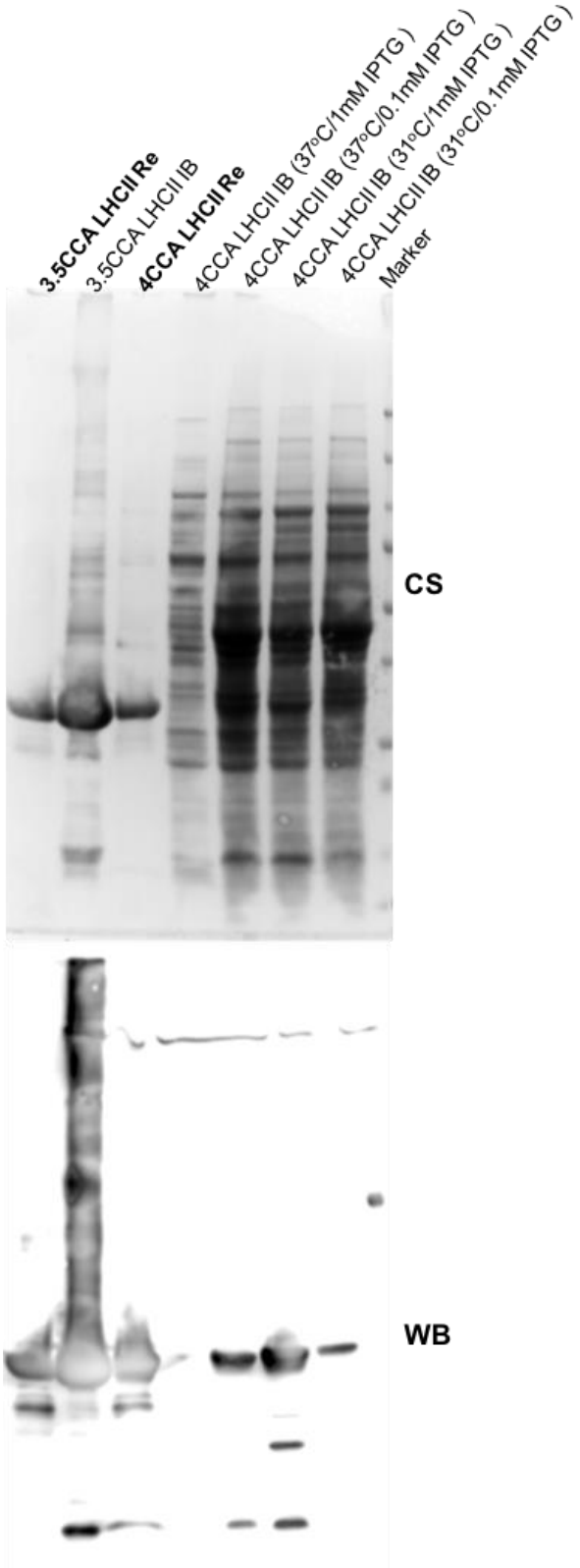


Figure S6.2. Original image of SDS-PAGE and Western blot. The lanes used in **Fig. 6.5EF** are highlighted in bold. CS – Coomassie Stain; WB – Western Blot. Expression tests with different inducing conditions were conducted on the 4CCA LHCII apo-protein.

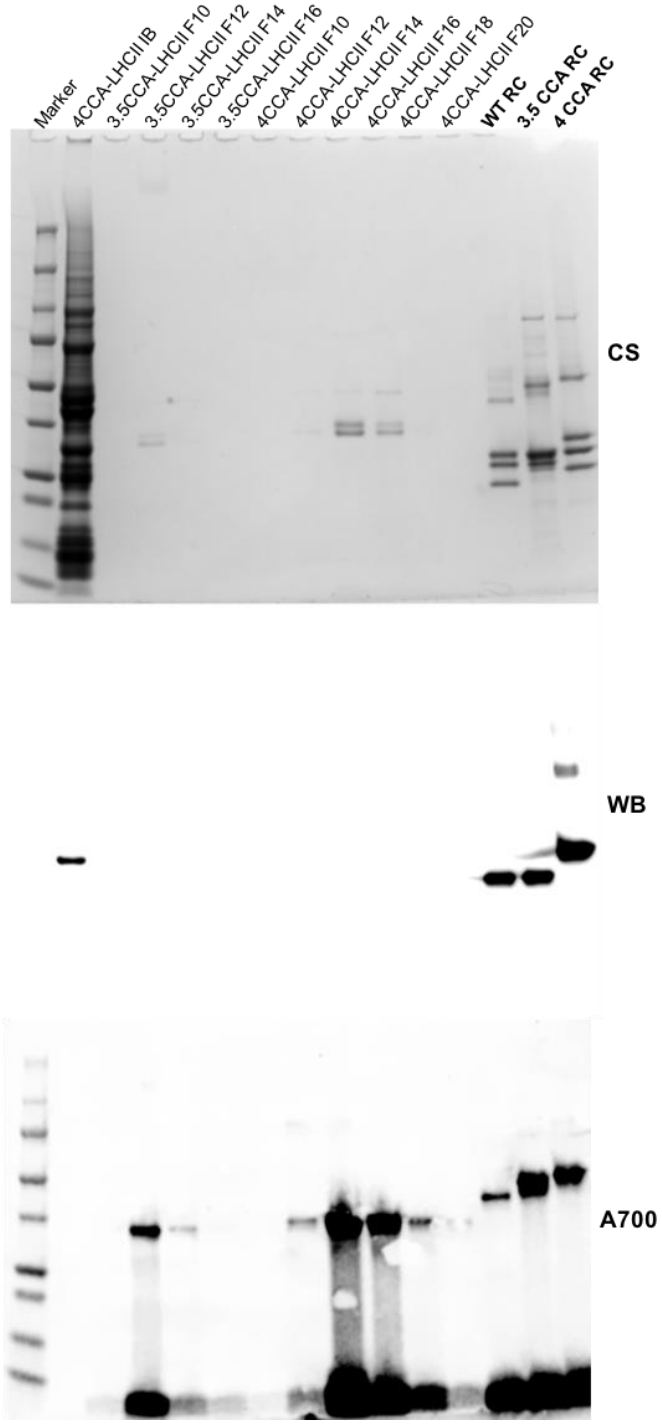


Figure S6.3. Original image of SDS-PAGE and Western blot. Samples were WT RC, 3.5CCA RC and 4CCA RC. Lanes used in **Fig. 6.5EF** are highlighted in bold. CS – Coomassie Stain; WB – Western Blot. The A700 channel was included to show the higher molecular weight band in WT RC, 3.5CCA RC and 4CCA RC were likely attributed to incomplete unfolding of RC complex since the band can be detected at absorbance around 700nm (A700) presumably due to associated pigments.

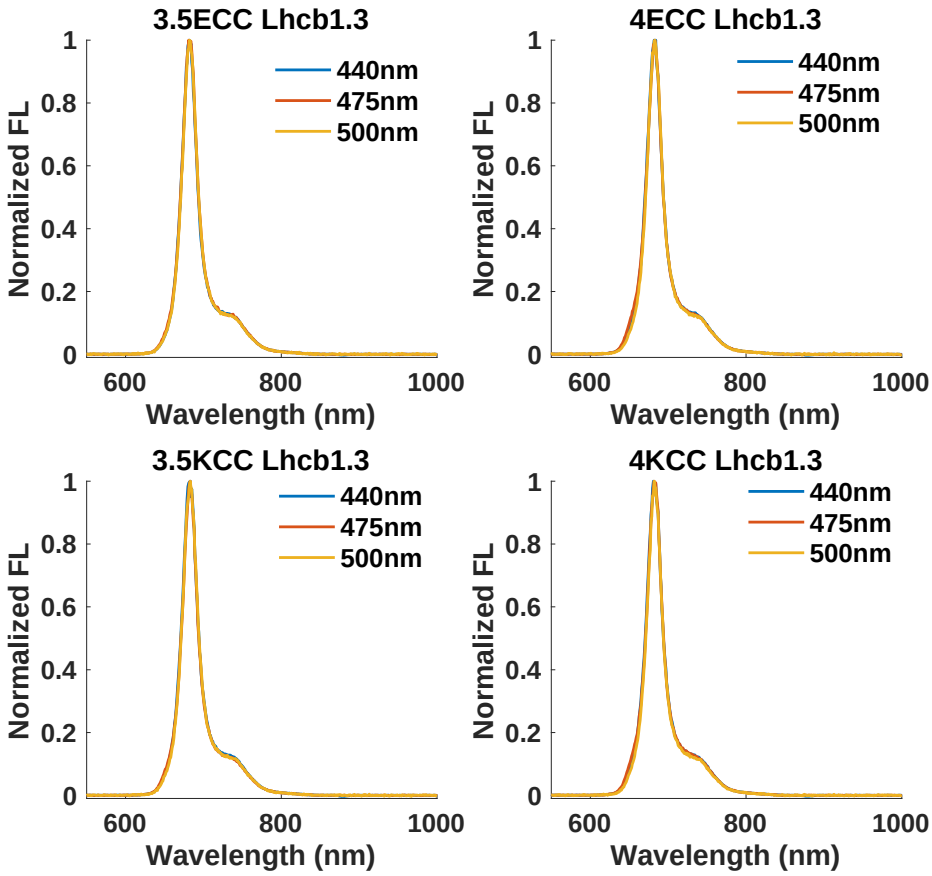


Figure S6.4. Normalized fluorescence spectra of all coiled-coil LHCII mutants. Normalized emission profiles of all purified LHCII coiled coil mutants are shown at excitation wavelength of 440 nm, 475 nm and 500 nm.

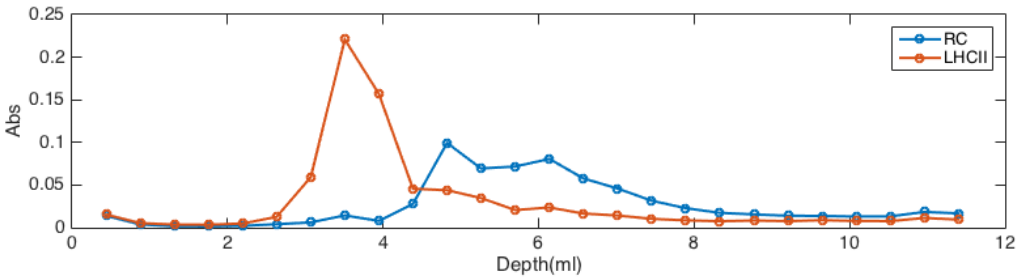


Figure S6.5. Absorbance profile of sucrose density gradient from Fig. 6.9A. Gradients was run with an amphipol stabilized 3.5K-CC RC/3.5E–CC LHCII mixture. Traces show absorbance at 804 nm for RC (blue) and 674 nm for LHCII (red). It was apparent that some LHCII was pulled to lower position and co-migrated with RCs.

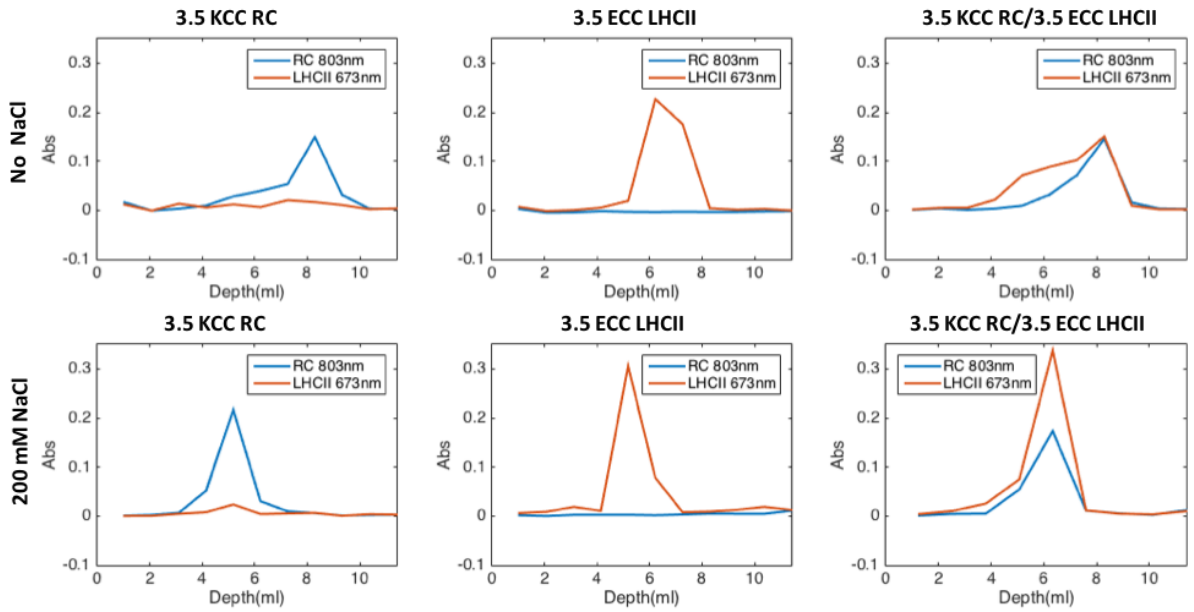


Figure S6.6. Absorbance profile of sucrose density gradients shown in Fig. 6.9B. Gradients were fractionated into 1 mL fractions. The top three profiles show the first three tubes in **Fig. 6.9B** (from left) and the lower three panels show the next three tubes. Traces showed the absorbance at 804 nm for RC (blue) and 674 nm for LHCII (red). It was apparent that some LHCII was pulled to a lower position and co-migrated with RCs.

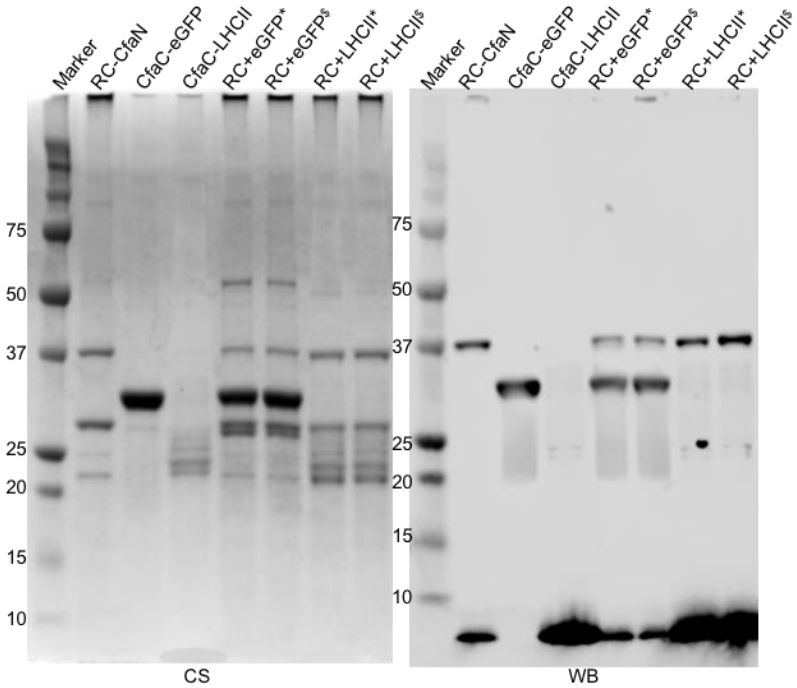


Figure S6.7. Use of Cfa intein – original gel used in Fig 6.11B. RC+eGFP or RC+LHCII indicated the reaction mixtures of Cfa mutants. HRP-conjugated anti-polyhistidine antibody

was used for detecting His-tags. The CfaC-LHCII was partially degraded. * - reaction buffer containing 5 mM TCEP; $- reaction buffer containing 10 mM DTT. CS – Coomassie stained; WB – Western blot.

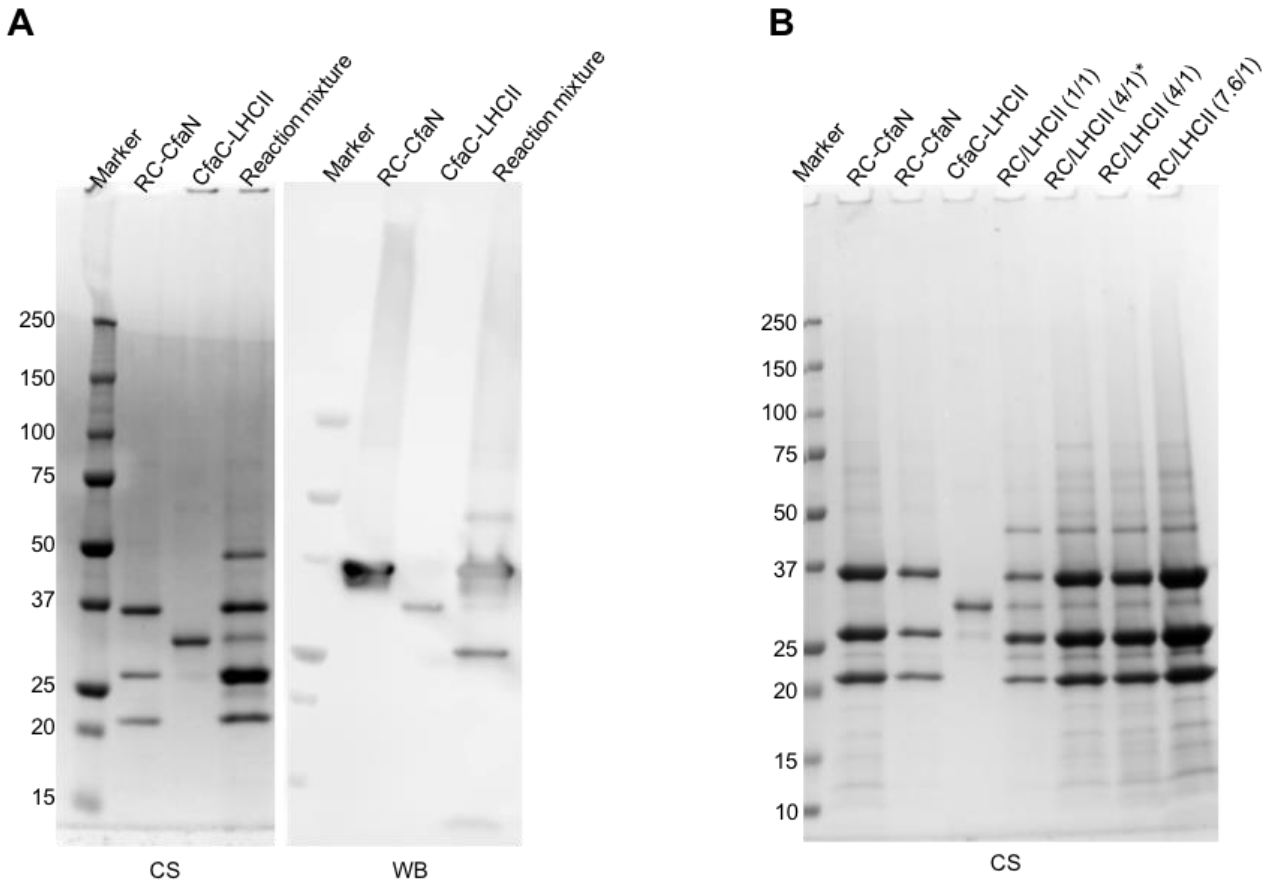


Figure S6.8. Use of Cfa intein – original gel used in Fig 6.11C. (A) Reaction mixture of RC-CfaN and CfaC-LHCII at a 1:1 ratio with controls of the two proteins alone. The formation of a PufM-Lhcb1.3 polypeptide conjugate was identified. (B) Variation of the ratio of RC/LHCII did not change the yield of PufM-Lhcb1.3 significantly. * - the total protein concentration was the same as the reaction mixture with RC/LHCII (1/1). All the other non-one ratios kept the concentrations of LHCs to 7.5 µM. CS – Coomassie stained; WB – Western blot.

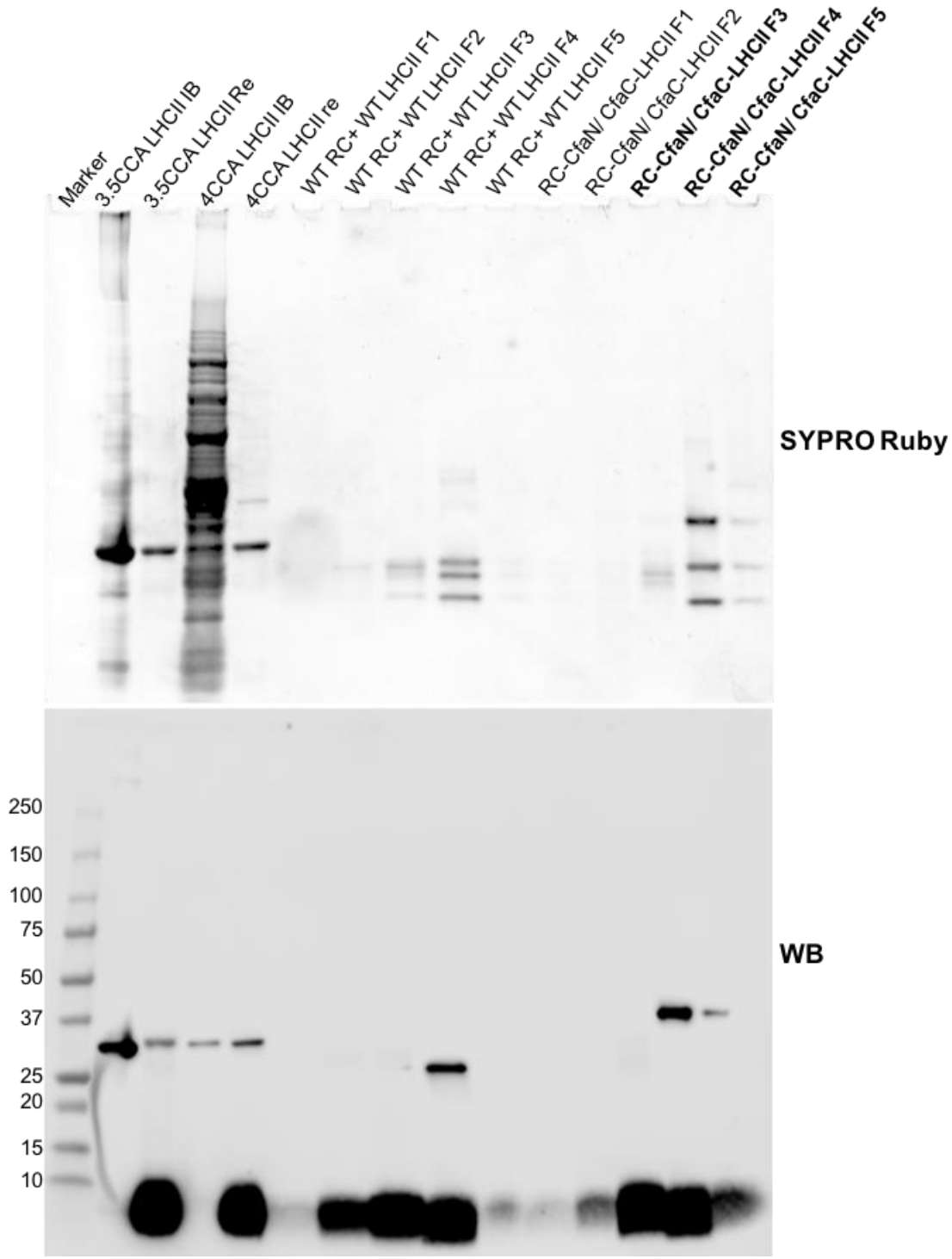


Figure S6.9. Use of Cfa intein – original gel used in Fig 6.11D. The sucrose density gradient in Fig. 6.11D was fractionated from the top into 1 mL fractions (F1 – F5). The F3 – F5 of the RC-CfaN+CfaC-LHCII gradient was equivalent to the B1 – B3 fractions described in the main text. The SDS-PAGE gel was stained with Sybro-Ruby (top panel) and western blotted with HRP-conjugated antiHis antibodies (bottom panel).

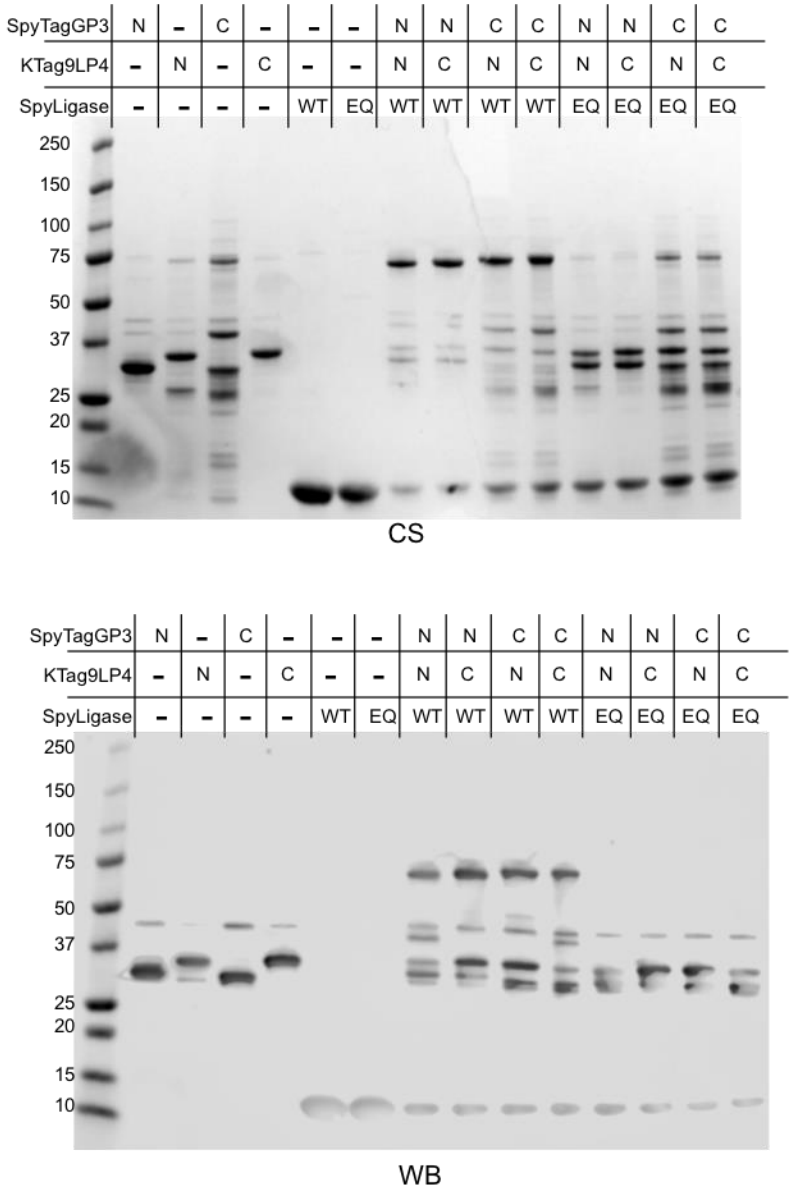


Figure S6.10. Use of SpyCoil – original gel used in Fig 6.15A. The top panel showed the Coomassie stained gel and the bottom panel was the result of anti-His western blot of the gel.

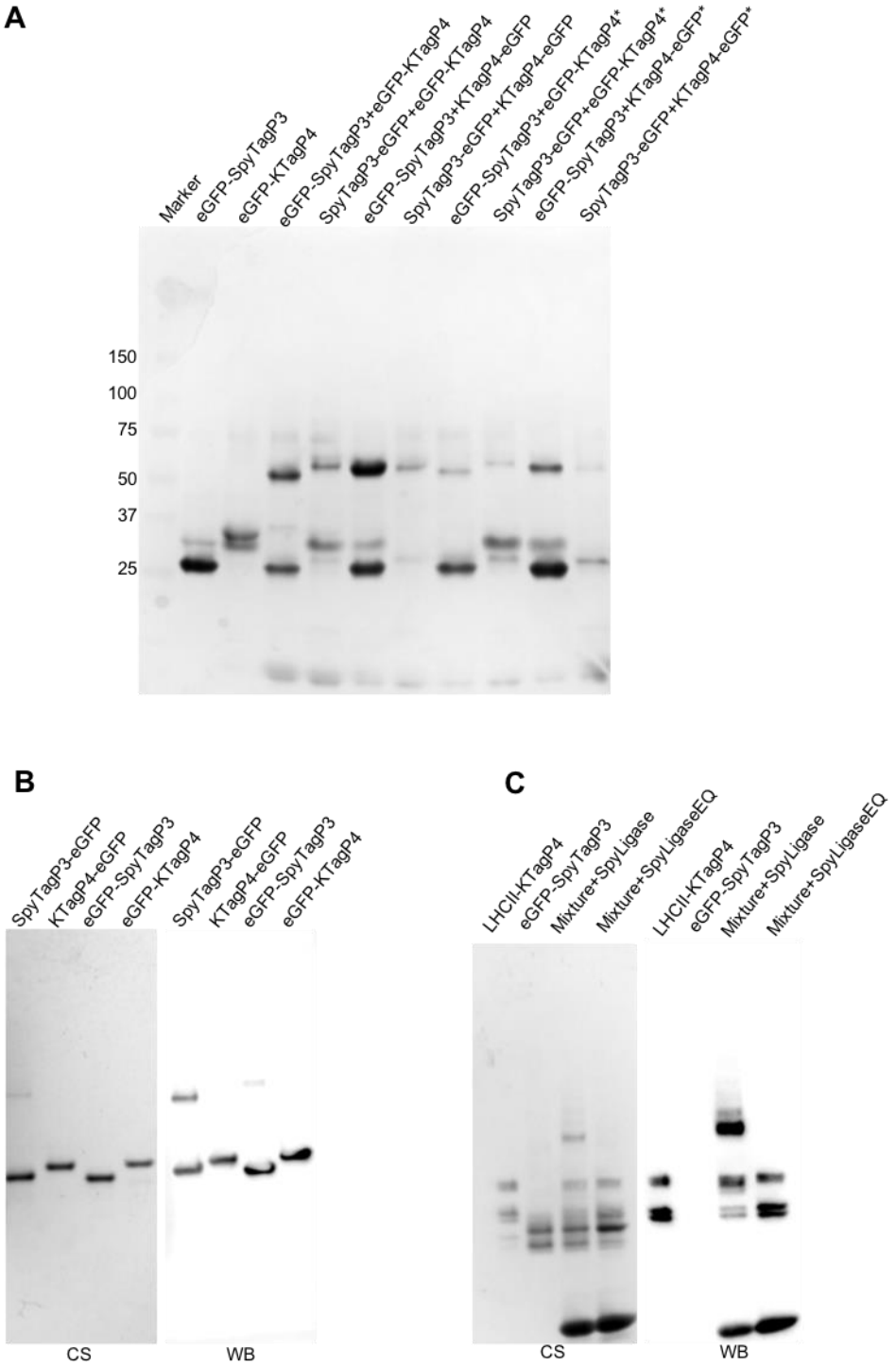


Figure S6.11. Use of SpyCoil – original gel used in Fig 6.15BC. (A) SDS-PAGE screening of SpyCoil-eGFP reactions displayed in **Fig. 6.15B**. Samples were in 20 mM Tris (pH 8.0), 500 mM Imidazole. * - reaction carried out with additional 500 mM NaCl. (B) SDS-PAGE/WB of purified SpyCoil-eGFP mutants. (C) SDS-PAGE/WB result of conjugation reaction between eGFP-SpyTagP3 and LHCII-KTagP4 in TrisDDM buffer shown in **Fig. 16C**. Mixture was a 1:2:10 combination of LHCII-KTagP4, eGFP-SpyTagP3 and SpyLigase. SpyLigaseEQ - Inactive mutant of SpyLigase; CS – Coomassie blue stained; WB – Western blot.

Table S6.1. Modifications of protein sequences.

Name	Modifications
AR1	EIAALKQEIAALKKENAALKWEIAALKQ -GGSGGTG-PufL/ EIAALKQEIAALKKENAALKWEIAALKQ -GGSGGTGGSG-PufLM –His_{10}
AR2	EIAAIKQEIAAIKKEIAAIKWEIAAIKQ -GGSGGTG-PufL / EIAAIKQEIAAIKKEIAAIKWEIAAIKQ - GGSGGTGGSG-PufM-His_{10}
AR3	EIAAIKQEIAAIKKEIAAIKWEIAAIKQ -GGSGGTG-PufL / EIAALKQEIAALKKENAALKWEIAALKQ - GGSGGTGGSG-PufM-His_{10}
AR4	EIAALKQEIAALKKENAALKWEIAALKQ -GGSGGTG-PufL / EIAAIKQEIAAIKKEIAAIKWEIAAIKQ - GGSGGTGGSG-PufM-His_{10}
3.5-CCA RC	LEQEIAALEKENAALEWEIAALEQ-GGSGGTG-PufL / PufM-His_{10}
3.5-CCB RC	lkqkiaalKyknaalkkkiaalkq-GGSGGTG- PufL / PufM-His_{10}
4-CCA RC	His_{10}-GGSGGTG-EIAALEQEIAALEKENAALEWEIAALEQ-ggsggtggsggtggsg-PufL
4-CCB RC	His_{10}-GGSGGTG-KiAAlkqkiaalKyknaalkkkiaalkq-ggsggtggsggtggsg-PufL
3.5-CCA LHCII	LEQEIAALEKENAALEWEIAALEQ-GGSGGTG-ΔLhcb1.3-His_6
3.5-CCB LHCII	lkqkiaalKyknaalkkkiaalkq-GGSGGTG-ΔLhcb1.3-His_6
4-CCA LHCII	EIAALEQEIAALEKENAALEWEIAALEQ-GGSGGTG-ΔLhcb1.3-His_6
4-CCB LHCII	KiAAlkqkiaalKyknaalkkkiaalkq-GGSGGTG-ΔLhcb1.3-His_6
RC-CfaN	PufL / PufM-GGSGGTG-CLSY(Cfa^N)dglp-SG-His_{10}
eGFP-CfaN	eGFP-SSSGGSGG-CLSY(Cfa^N)dglp-SG-His_6
CfaC-eGFP	His_6-SSG-VKIISRKSLGTQNVYDIGVEKDHNFLLKNGLVASNCFN-SGGSGGS-eGFP
CfaC-LHCII	VKIISRKSLGTQNVYDIGVEKDHNFLLKNGLVASNCFN-GGSGGTG-Lhcb1.3-His_6
SpyTagP3-eGFP	His_6-SSG-LVPRGS-HMG-AHIVMVDAYKPTK-G-KIQALQQKIKQLKQKIAALK-SGGSGGS-eGFP
KTagP4-eGFP	His_6-SSG-LVPRGS-HMG-ATHIKFSKRD-GSGGSGGSG-QIAALEQEIAALEQEIAALE-SGGSGGS-eGFP
eGFP-SpyTagP3	eGFP-GSGSGSG-AHIVMVDAYKPTK-G-KIQALQQKIKQLKQKIAALK-SGSSSLE-His_6
eGFP-KTagP4	eGFP-GSGSGSG-ATHIKFSKRD-GSGGSGGSG-QIAALEQEIAALEQEIAALE-SGSSSLE-His_6
LHCII-KTagP4	Lhcb1.3-GSGDGTG-ATHIKFSKRD-GSGGSGGSG-QIAALEQEIAALEQEIAALE-SSG-His_6

The linker was labelled with underscore to distinguish from other functional motifs. The full sequence of the N-terminal domain of the split intein Cfa is not shown. Reactive cysteine of Cfa intein was highlighted in red/bold. Δ indicated the N-terminal 14 amino acids (-GSRKTVAKPKGPSG-) after first methionine was deleted to shorten the separation of the final conjugate products.

Chapter 7

7. Summary and Outlook

As outlined in Section 1.8 of Chapter 1, the concept underlying this project was to engineer hybrid photosystems in which the solar energy conversion properties of the purple bacterial RC are expanded by directly interfacing them with LHCs from plants. The motivation was to address some of the limitations displayed by purple bacterial RCs and RC-LH1 complexes when incorporated into photoelectrochemical devices for solar energy conversion, energy storage and sensing, in particular the relatively weak response to visible light. The main idea was to explore ways to genetically-modify both purple bacterial RCs and plant LHCs to form single macromolecular complexes, adopting a concept from synthetic biology of treating the proteins as components that can be interfaced in a predictable way. This was successfully achieved using the SpyCatcher/SpyTag protein domain as the interface, a main attribute of the approach being the highly stable way in which the interface domain connects the RC and LHC components through covalent locking. Two other interface domains were also explored. Despite these to some extent work with water-soluble test proteins, they were less successful with the RC and LHC integral membrane proteins. A third alternative interface domain, the SpyCoil, was developed and successfully tested including with LHCII, but unfortunately there was insufficient time to attempt using SpyCoil to form a chimera between LHCII and the RC.

Alongside this work on all-protein systems, conjugates between RCs, LHCIIs and CdTe QDs were also assembled and characterised. The initiative is to target the poor spectral overlap between plant LHC and RC to enhance the ET efficiency with the least modification of photo-components. At the beginning of the project a serial of researches had been published on interfacing His-tagged trimeric LHCII or non-His-tagged *Rba. sphaeroides* RCs with QDs, and so it seemed feasible to use QDs to build hybrids between RCs and LHCs. To bring control to the system it was decided to use the same, ten-residue, tight binding His-tag on both the RC and LHCII, and to select

QDs that would bind multiple copies of protein and also, due to their size and material, provide an energy bridge between the Q_y bands of LHCII and the RC. As the project developed and data on conjugate architecture, absorbance and emission properties were collected it became clear that the clear understanding and precise control of energy coupling depended on thermodynamics of the system created.

The account in Chapter 3 explains how it was established that water-soluble CdTe QDs could form a binding interaction with His_{10}-tag modified *Rba. sphaeroides* RCs. Use of the tag gave a geometric uniformity and specificity to the RC-QD interaction, a marked improvement on poorly-characterised electrostatic or hydrophobic interactions reported in previous studies of binding of RCs to water-soluble and water-insoluble QDs. The tag is presumed to have uniformly oriented the RCs with the P870 primary electron donor closest to the QD surface. The dissociation constant between a single RC and a QD was estimated as being 8.1 nM, indicating a tight binding interaction, and this interaction could be well explained by a model assuming multiple, independent binding events. Bound RCs were found to quench QD fluorescence and, conversely, QD excitation was found to drive photochemistry in WT or carotenoid-less RCs and drive RC emission in photochemically-inactive RCs. The quenching of QD emission was sensitive to the extent of spectral overlap with the RC absorbance, consistent with a FRET mechanism for the energy transfer. The estimated FRET distance *R* was consistent with morphologies of the RC-QD conjugates predicted from modelling of the RC/QD interaction and measured through TEM imaging and DLS. The single donor/acceptor FRET efficiency, E_{DA}, was of the order of 53% for the smallest conjugates involving WT_H RCs, and somewhat higher (55-60%) for pigment-replacement mutant RCs with enhanced spectral overlap. Evidence suggestive of exciton-exciton annihilation was also seen when photochemically-inactive RCs were packed around a central QD at a high RC:QD ratio.

The data and analysis outlined in Chapter 3 was encouraging in that it showed that it was possible to build RC/QD conjugates in a predictable way and characterise their energy transfer properties, unveiling the possibility to use the QD as a synthetic antenna for the RC, capturing light across the visible and UV regions and delivering it to the RC via the 750 nm emission band. Accordingly, it was decided to attempt to

build tri-component conjugates in which a mixed shell of RCs and plant LHCs would surround the central QD assembly hub.

As explained in Chapter 4, the first step in this was to build and characterise conjugates between QDs and His_{10}-tagged LHCII. As with the research going on in parallel with chimeras this required mastering the approach first developed more than 20 years ago of reconstituting LHCII from proteins expressed in *E. coli* and purified pigments with the help and advice from Prof. Roberta Croce from the Vrije Universiteit Amsterdam, who has used this procedure for many years to study LHCII and LHCI. Based on the high quality of authentic refold holo-LHCs, it was then possible to investigate LHCII/QD conjugates of various sizes and show energy transfer from LHCII to the QD assembly hub. Importantly, when the thermodynamics of assembly were investigated it was found that the dissociation constant between a single LHCII and a QD was 8.2 nM, almost identical to that for the RC with the same size of Histag. This suggested that it should be possible to build tri-component conjugates in a predictable and adjustable way. As explained in Chapter 4, this proved to be the case.

The large part of Chapter 4 describes the natures and properties of these tricomponent conjugates. The thermodynamic process of self-assembly could be well explained by a scheme of competitive binding to the QD by the RC and LHCII. Architectures of the resulting conjugates consistent with theoretical physical dimensions of the RC and somewhat smaller LHCII. It was established that, loose packing up to a density of at least 10 proteins per QD, the average numbers of RC and LHCII were determined by their concentrations in solution, providing a simple means to control the average composition of the conjugate. Together, direct and indirect FRET produced an efficient energy flow from the LHCII energy harvesters to the RC photochemical traps in these conjugates. The fluxes of energy flow along the two routes could be tuned in a predictable manner by simply varying the stoichiometry of $LHCII_H$ to RC_H in the mix used for self-assembly, and by varying the total amount of protein per QD it was possible to tune the overall efficiency of energy transfer from nearly zero to about 60%. The experiments and analysis described in Chapter 4 show that that is was possible to understand both the thermodynamics of assembly and energy transfer characteristics of the RC/LHCII/QD conjugates, and it should be the case that this understanding can be readily transferred to other types of

protein/nanocrystal conjugate. The approach described in the chapter also sets out a blueprint for the future self-assembly and analysis of more complex modular bio-conjugates employing a wider range of light harvesting and electron transfer components based on a systematic approach, which is an important concept of synthetic biology.

Chapters 5 and 6 of this thesis set out the main findings from attempts to build all-protein hybrid photosystems. All of the peptidyl linking strategies outlined in Chapter 6 could, in principle, have directed the self-assembly of the desired protein macromolecule *in vitro* and, because of the peptidyl nature of these components, they could potentially be adapted for use *in vivo* without massive modification or intervention of the interested biological systems. A possible benefit of this is that the design could be used with species other than model organisms such as *Rba. sphaeroides* or *E. coli*. In the event attempts to use coiled-coils and split inteins proved limited successes (only a summary of key results is reported in Chapter 6), exposing a need to substantially optimize their function in accordance with the purpose of their use at very different context to stand-alone. A more successful experience was had with SpyCoil, which used the specificity of a coiled coil to improve the functionality of the SpyTag/KTag/SpyLigase system. This gave promising results including with LHCII, but there was insufficient time at the end of the project to test SpyCoil's ability to self-assemble RC-LHCII chimeras. Although this may be possible in future work some caution must be expressed as an experience of this project and related work has been that the *Rba. sphaeroides* does not always tolerate the addition of extra components to the PufL and PufM proteins of the RC. This requires a certain amount of trial and error in construct design and, for example, is the reason that the linking strategy described in Chapter 6 was to connect SpyCatcher-modified RCs with SpyTag-LHCIIs. The reverse strategy could not be tested because the small SpyTag peptide is not tolerated by the RC even though the much larger SpyCatcher domain is.

As described in Chapter 5, despite these issues the SpyCatcher/SpyTag protein fusion emerged as the most versatile and reliable approach to construct large multi-protein complexes due to its ultra-stability, compatibility and specific isopeptide chemistry (and especially its attribute of having no obvious side reactions). The work described in Chapter 5 showed that genetically-adapting two types of evolutionarily-

disparate photosynthetic membrane proteins with this extramembrane interface domain enabled self-assembly of a chimeric photosystem in which UV/near-IR solar energy conversion by the *Rba. sphaeroides* RC was augmented by visible light capture by the Chls of LHCI and LHCII with complementary absorption profiles. These chimeras could be formed with a high yield by simply mixing adapted proteins in solution and were stabilised by the isopeptide bond within SpyTag/SpyCatcher that covalently linked the LHC protein to one of the component proteins of the RC, locking the structure. The resulting chimeras displayed energy Chl-to-BChl energy transfer in solution with efficiencies to about 20% for single RC-LHC pair and tethering the LHC to either one or two RCs did not interfere with the RCs ability to carry out charge separation in solution or use this charge separation to drive a photocurrent when adhered to an electrode. Although it turned out that Chl-to-BChl energy transfer could also be detected, with a roughly similar efficiency, when unadapted RCs and LHCIIs were co-adsorbed on an electrode, the chimera approach meant that energy transfer was guaranteed in wide range of contents (e.g. in solution). For Some instances of photoelectrochemical cells are based on solutions of photoproteins rather than films on electrode surfaces, and hence it will be particularly interesting to see the application based on these bottom-up designed photosystems.

In summary, these approaches depending on the fundamental concept of synthetic biology of using genetic encoding to adapt naturally incompatible components to interface in a predictable way, either directly or indirectly via a complementary synthetic component, created novel macromolecular photosystems that were predictable and programmable. In addition to providing novel photosynthetic structures and energy transfer pathways to explore, these polychromatic photosystems provide interesting new materials for biohybrid devices that in recent years have expanded in application beyond photoelectrochemical solar energy conversion to fuel molecule synthesis, energy storage, biosensing, touch sensing and photodetection. In addition, the demonstrated flexibility with which RCs and LHCs could be interfaced opens the possibility of constructing more elaborate, self-assembling chimeric photosystems that employ multiple orthogonal linking modules and a wider range of photosynthetic and redox proteins that, despite being separated by billions of years of evolution, can be adapted for future solar energy conversion through genetic programming of predictable interfaces or artificial material that are

chemical compatible. Finally, this approach should be compatible *in vivo* without massive fundamental re-engineering of host organisms, providing a route that could be taken to achieve bottom up redesign of photosynthetic membranes *in situ*. With the power of directed evolution and rational computational design it seems certain that there will be creation in the future of new potent linking strategies and, together with a deeper comprehension of both natural and artificial solar energy conversion, the era of bottom-up redesign of photosynthesis will bloom.

The bottom-up construction of chimeric photosystem focusing on the initial light-electron conversion is a novel approach to improve photosynthesis. This work could pave the way toward systematic redesign of photosynthetic membrane configuration, opening the opportunity to re-wire energy transfer pathway for better light utility. Future work should emphasize on the *in vivo* realisation of the design but did not have to restricted to the selected components showed in this work (plant LHCs and bacterial RCs). The redesign could be a wider range of compatible photosynthetic complexes (e.g. proteorhodopsin or phycobilin). Some challenges need to be resolved before seeing real improvement on photosynthesis particularly the tolerance of host organism to the sustainable modification. Natural photosynthetic membrane had a very elegant core assemblies with pigment forming a defined energy transfer pathway (e.g. efficient energy transfer in plant $C_2S_2M_2$ either directly from major LHCII to PSII or mediated via minor complex CP24 or CP26), while the organization of peripheral harvesting antennas just need to be compacted surrounding the core units. This nuance in the requirement of architecture precision deserve more attention for achieving an optimal balance between design complexity and effectiveness of design. For example, QDs mediated self-assemble process resulted in a heterogeneous mixture of conjugates that might seem less controllable than SpyTag/SpyCatcher approach, however, it did not require massive modification of protein sequence and hence would be a better strategy to link RC core units to a swarm of light harvesting antennas. The toxic Cd/Te QD could be replace by more biocompatible assembly hub such as carbon dots since the thermodynamics described in this thesis is a general description of binding despite the components of choices. Once a good interface can be established between photosynthetic complexes and carbon dots, we should be able to remodel the photosynthetic membrane assemblies either *in vitro* and more importantly *in vivo*.

Bibliography

1. Drive, M. R. & Hill, C. From volcanic origins of chemoautotrophic life to Bacteria , Archaea and Eukarya. *Phil. Trans. R. Soc. B* **361,** 1787–1808 (2006).
2. Alexandratos, N. & Bruinsma, J. World agriculture: towards 2015/2030: an FAO perspective. *Land Use Policy* **20,** 1–10 (2012).
3. Erb, T. J. & Zarzycki, J. Biochemical and synthetic biology approaches to improve photosynthetic CO2 -fixation. *Curr. Opin. Chem. Biol.* **34,** 72–79 (2016).
4. Long, S. P., Marshall-Colon, A. & Zhu, X. Review Meeting the Global Food Demand of the Future by Engineering Crop Photosynthesis and Yield Potential. *Cell* **161,** 56–66 (2015).
5. Becker, E. W. Micro-algae as a source of protein. *Biotechnol. Adv.* **25,** 207–210 (2007).
6. Ringsmuth, A. K., Landsberg, M. J. & Hankamer, B. Can photosynthesis enable a global transition from fossil fuels to solar fuels , to mitigate climate change and fuel-supply limitations ? *Renew. Sustain. Energy Rev.* **62,** 134–163 (2016).
7. Dudley, B. *BP Statistical Review of World Energy 2018*. (2018).
8. Al-Ghussain, L. Global Warming : Review on Driving Forces and Mitigation. *Environ. Prog. Sustain. Energy* **38,** 13–21 (2019).
9. Kang, S. & Eltahir, E. A. B. North China Plain threatened by deadly heatwaves due to climate change and irrigation. *Nat. Commun.* **9,** 1–9 (2018).
10. IPCC. *Climate Change 2014: Synthesis Report. Climate Change 2014: Synthesis* (2014). doi:10.1256/004316502320517344
11. Mann, M. E. *et al.* Influence of Anthropogenic Climate Change on Planetary Wave Resonance and Extreme Weather Events. *Sci. Rep.* **7,** 1–10 (2017).
12. Lewis, N. S. & Crabtree, G. *Basic Research Needs For Solar Energy Utilizaton*. (2005).
13. Nocera, D. G. & Nash, M. P. Powering the planet: Chemical challenges in solar energy utilization. *Proc Natl Acad Sci U S A* **104,** 15729–15735 (2007).
14. The European Economic and Social Committee. *Energy Roadmap 2050*. *European Commission* (2011).
15. Department for Business Energy & Industrial Strategy. *UK Energy in Brief 2018*. (2018).
16. Department for Business Energy & Industrial Strategy. *UK Energy in Brief 2017*. *Department for Business Energy and Industrial Strategy* (2017).
17. Strategy, D. for B. E. & I. *UK Energy in Brief 2016*. (2016).
18. Mai, T., Sandor, D., Wiser, R. & Schneider, T. *Renewable Electricity Futures Study: Executive Summary*. **1,** (2012).
19. Mi, Z. *et al.* Chinese CO_2 emission flows have reversed since the global financial crisis. *Nat. Commun.* **8,** 1–10 (2017).
20. National Development and Reform Commission. *China 13th Renewable Energy Development Five Year Plan (2016-2020)*. (2016).
21. Blankenship, R. E. in *Molecular Mechanisms of Photosynthesis* (ed. Blankenship, R. E.) 26–41 (Wiley, 2008). doi:10.1002/9780470758472.ch1
22. Buick, R. When did oxygenic photosynthesis evolve? *Philos. Trans. R. Soc. B Biol. Sci.* **363,** 2731–2743 (2008).
23. Reeburgh, W. S. Figures Summarizing the Global Cycles of Biogeochemically Important Elements. *Bullentin Ecol. Soc. Am.* **78,** 260–267 (1997).
24. Bryant, D. A. *et al. Candidatus* Chloracidobacterium thermophilum: An Aerobic Phototrophic Acidobacterium. *Science* **272,** 523–527 (2007).
25. Zeng, Y., Feng, F., Medová, H., Dean, J. & Koblí, M. Functional type 2 photosynthetic reaction centers found in the rare bacterial phylum Gemmatimonadetes. *Proc Natl Acad Sci U S A* **111,** 7795–7800 (2014).
26. Cardona, T. A fresh look at the evolution and diversification of photochemical reaction centers. *Photosynth. Res.* **126,** 111–134 (2015).
27. Orf, G. S., Gisriel, C. & Redding, K. E. Evolution of photosynthetic reaction centers :

insights from the structure of the heliobacterial reaction center. *Photosynth. Res.* **138,** 11–37 (2018).
28. Knoll, A. H. & Nowak, M. A. The timetable of evolution. *Sci. Adv.* **3,** e1603076 (2017).
29. Berner, R. A. Perspective Atmospheric oxygen over Phanerozoic time. *Proc Natl Acad Sci U S A* **96,** 10955–10957 (1999).
30. Pisciotta, J. M., Zou, Y. & Baskakov, I. V. Light-Dependent Electrogenic Activity of Cyanobacteria. *PLoS One* **5,** e10821 (2010).
31. Björn, L. O., Papageorgiou, G. C., Blankenship, R. E. & Govindjee. A viewpoint: Why chlorophyll *a*? *Photosynth. Res.* **99,** 85–98 (2009).
32. Murakami, A., Miyashita, H., Iseki, M., Adachi, K. & Mimuro, M. Chlorophyll <i>d<\i> in an Epiphytic Cyanobacterium of Red Algae. *Science* **303,** 1633 (2004).
33. Nürnberg, D. J. *et al.* Photochemistry beyond the red limit in chlorophyll *f*-containing photosystems. *Science* **360,** 1210–1213 (2018).
34. Stomp, M., Huisman, J., Stal, L. J. & Matthijs, H. C. P. Colorful niches of phototrophic microorganisms shaped by vibrations of the water molecule. *ISME J.* **1,** 271–282 (2007).
35. Loughlin, P., Lin, Y. & Chen, M. Chlorophyll *d* and *Acaryochloris marina*: current status. *Photosynth. Res.* **116,** 277–293 (2013).
36. Ho, M., Soulier, N. T., Canniffe, D. P., Shen, G. & Bryant, D. A. Light regulation of pigment and photosystem biosynthesis in cyanobacteria. *Curr. Opin. Plant Biol.* **37,** 24–33 (2017).
37. Scholes, G. D. *et al.* Lessons from chlorophylls: Modifications of porphyrinoids towards optimized solar energy conversion. *Molecules* **19,** 15938–15954 (2014).
38. Jones, M. R. The petite purple photosynthetic powerpack. *Biochem. Soc. Trans.* **37,** 400–407 (2009).
39. Caffarri, S., Tibiletti, T., Jennings, R. C. & Santabarbara, S. A Comparison Between Plant Photosystem I and Photosystem II Architecture and Functioning. *Curr. Protein Pept. Sci.* **15,** 296–331 (2014).
40. Nelson, N. & Junge, W. Structure and Energy Transfer in Photosystems of Oxygenic Photosynthesis. *Annu. Rev. Biochem.* **84,** 659–683 (2015).
41. Barber, J. Photosystem II: the water splitting enzyme of photosynthesis and the origin of oxygen in our atmosphere. *Q. Rev. Biophys.* **49,** e14 (2016).
42. Umena, Y., Kawakami, K., Shen, J. & Kamiya, N. Crystal structure of oxygen-evolving Photosystem II at a resolution of 1.9 Å. *Nature* **473,** 55–60 (2011).
43. Kargul, J., Janna Olmos, J. D. & Krupnik, T. Structure and function of Photosystem I and its application in biomimetic solar-to-fuel systems. *J. Plant Physiol.* **169,** 1639–1653 (2012).
44. Ben-Shem, A., Frolow, F. & Nelson, N. Crystal structure of plant Photosystem I. *Nature* **426,** 630–635 (2003).
45. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. Structure of the protein subunits in the photosynthetic reaction centre of *Rhodopseudomonas viridis* at 3Å resolution. *Nature* **318,** 618–624 (1985).
46. Allen, J. P., Feher, G., Yeates, T. O., Komiya, H. & Rees, D. C. Structure of the reaction center from *Rhodobacter sphaeroides* R-26: the cofactors. *Proc. Natl. Acad. Sci.* **84,** 5730–5734 (1987).
47. Allen, J. P., Feher, G., Yeates, T. O., Komiya, H. & Rees, D. C. Structure of the reaction center from *Rhodobacter sphaeroides* R-26: the protein subunits. *Proc. Natl. Acad. Sci.* **84,** 6162–6166 (1987).
48. Nogi, T., Fathir, I., Kobayashi, M., Nozawa, T. & Miki, K. Crystal structures of photosynthetic reaction center and high-potential iron-sulfur protein from *Thermochromatium tepidum*: Thermostability and electron transfer. *Proc. Natl. Acad. Sci.* **97,** 13561–13566 (2000).
49. Gisriel, C. *et al.* Structure of a symmetric photosynthetic reaction center-photosystem. *Science* **357,** 1021–1025 (2017).
50. Feher, G., Allen, J. P., Okamura, M. Y. & Rees, D. C. Structue and function of

bacterial photosynthetic reaction centers. *Nature* **339,** 111–116 (1989).
51. Zinth, W. & Wachtveitl, J. The first picoseconds in bacterial photosynthesis - Ultrafast electron transfer for the efficient conversion of light energy. *ChemPhysChem* **6,** 871–880 (2005).
52. Hoff, A. J. & Deisenhofer, J. Photophysics of photosynthesis. Structure and spectroscopy reaction centers of purple bacteria. *Phys. Rep.* **287,** 1–247 (1997).
53. Heathcote, P. & Jones, M. R. in *Comprehensive Biophysics* (eds. Egelman, E. H. & Ferguson, S.) **8,** 115–144 (Academic Press, 2012).
54. Ermler, U., Fritzsch, G., Buchanan, S. K. & Michel, H. Structure of the photosynthetic reaction centre from *Rhodobacter sphaeroides* at 2.65 Å resolution : cofactors and protein-cofactor interactions. *Structure* **2,** 925–936 (1994).
55. Cardona, T., Sedoud, A., Cox, N. & Rutherford, A. W. Charge separation in Photosystem II : A comparative and evolutionary overview. *BBA - Bioenerg.* **1817,** 26–43 (2012).
56. Malferrari, M. *et al.* Trehalose matrix effects on charge-recombination kinetics in Photosystem I of oxygenic photosynthesis at different dehydration levels. *BBA - Bioenerg.* **1857,** 1440–1454 (2016).
57. Feher, G., Allen, J. P., Okamura, M. Y. & Rees, D. C. Structure and function of bacterial photosynthetic reaction centres. *Nature* **339,** 111–116 (1989).
58. Kleinfeld, D., Okamura, M. Y. & Feher, G. Electron-transfer kinetics in photosynthetic reaction centers cooled to cryogenic temperatures in the charge-separated state: evidence for light-induced structural changes. *Biochemistry* **23,** 5780–5786 (1984).
59. Kleinfeld, D., Okamura, M. Y. & Feher, G. Electron transfer in reaction centers of *Rhodopseudomonas sphaeroides*. II. Free energy and kinetic relations between the acceptor states $Q_A^-Q_B^-$ and $Q_AQ^{2-}{}_B$. *Biochim. Biophys. Acta - Bioenerg.* **809,** 291–310 (1985).
60. King, B. A., McAnaney, T. B., DeWinter, A. & Boxer, S. G. Excited State Energy Transfer Pathways in Photosynthetic Reaction Centers. 3. Ultrafast Emission from the Monomeric Bacteriochlorophylls. *J. Phys. Chem. B* **104,** 8895–8902 (2000).
61. van Brederode, M. E. & van Grondelle, R. New and unexpected routes for ultrafast electron transfer in photosynthetic reaction centers. *FEBS Letters* **455,** 1–7 (1999).
62. Jansson, S. The light-harvesting chlorophyll *a b*-binding proteins. *BBA - Bioenerg.* **1184,** 1–19 (1994).
63. Jansson, S. A guide to the Lhc genes and their relatives in *Arabidopsis*. *Trends Plant Sci.* **4,** 236–240 (1999).
64. Liu, Z. *et al.* Crystal structure of spinach major light-harvesting complex at 2.72 Å resolution. *Nature* **428,** 287–292 (2004).
65. Standfuss, J., Van Scheltinga, A. C. T., Lamborghini, M. & Kühlbrandt, W. Mechanisms of photoprotection and nonphotochemical quenching in pea light-harvesting complex at 2.5 Å resolution. *EMBO J.* **24,** 919–928 (2005).
66. Su, X., Wei, X., Zhu, D., Chang, W. & Liu, Z. Structure and assembly mechanism of plant $C_2S_2M_2$-type PSII-LHCII supercomplex. *Science* **820,** 815–820 (2017).
67. Krüger, T. P. J., Novoderezhkin, V. I., Ilioaia, C. & van Grondelle, R. Fluorescence spectral dynamics of single LHCII trimers. *Biophys. J.* **98,** 3093–3101 (2010).
68. van Oort, B., van Hoek, A., Ruban, A. V. & van Amerongen, H. Aggregation of Light-Harvesting Complex II leads to formation of efficient excitation energy traps in monomeric and trimeric complexes. *FEBS Lett.* **581,** 3528–3532 (2007).
69. Betterle, N., Ballottari, M., Hienerwadel, R., Dall'Osto, L. & Bassi, R. Dynamics of zeaxanthin binding to the Photosystem II monomeric antenna protein Lhcb6 (CP24) and modulation of its photoprotection properties. *Arch. Biochem. Biophys.* **504,** 67–77 (2010).
70. Miloslavina, Y. *et al.* Far-red fluorescence: A direct spectroscopic marker for LHCII oligomer formation in non-photochemical quenching. *FEBS Lett.* **582,** 3625–3631 (2008).
71. Caffarri, S., Kouřil, R., Kereïche, S., Boekema, E. J. & Croce, R. Functional

architecture of higher plant Photosystem II supercomplexes. *EMBO J.* **28,** 3052–3063 (2009).
72. Natali, A., Roy, L. M. & Croce, R. *In Vitro* Reconstitution of Light-harvesting Complexes of Plants and Green Algae. *J. Vis. Exp.* **92,** 1–13 (2014).
73. Croce, R., Morosinotto, T., Castelletti, S., Breton, J. & Bassi, R. The Lhca antenna complexes of higher plants Photosystem I. *Biochim. Biophys. Acta - Bioenerg.* **1556,** 29–40 (2002).
74. Wientjes, E. & Croce, R. The light-harvesting complexes of higher-plant Photosystem I: Lhca1/4 and Lhca2/3 form two red-emitting heterodimers. *Biochem. J.* **433,** 477–485 (2011).
75. Qin, X., Suga, M., Kuang, T. & Shen, J.-R. Structural basis for energy transfer pathways in the plant PSI-LHCI supercomplex. *Science* **348,** 989–995 (2015).
76. Krüger, T. P. J., Wientjes, E., Croce, R. & van Grondelle, R. Conformational switching explains the intrinsic multifunctionality of plant light-harvesting complexes. *Proc. Natl. Acad. Sci.* **108,** 13516–13521 (2011).
77. Prince, S. M. *et al.* Apoprotein structure in the LH2 complex from *Rhodopseudomonas acidophila* strain 10050: modular assembly and protein pigment interactions11Edited by R. Huber. *J. Mol. Biol.* **268,** 412–423 (1997).
78. Koepke, J., Hu, X., Muenke, C., Schulten, K. & Michel, H. The crystal structure of the light-harvesting complex II (B800–B850) from *Rhodospirillum molischianum*. *Structure* **4,** 581–597 (1996).
79. Yu, L.-J., Suga, M., Wang-Otomo, Z.-Y. & Shen, J.-R. Structure of photosynthetic LH1–RC supercomplex at 1.9 Å resolution. *Nature* **556,** 209–213 (2018).
80. Qian, P., Siebert, C. A., Wang, P., Canniffe, D. P. & Hunter, C. N. Cryo-EM structure of the *Blastochloris viridis* LH1–RC complex at 2.9 Å. *Nature* **556,** 203–208 (2018).
81. Qian, P. *et al.* Three-Dimensional Structure of the *Rhodobacter sphaeroides* RC-LH1-PufX Complex: Dimerization and Quinone Channels Promoted by PufX. *Biochemistry* **52,** 7575–7585 (2013).
82. Papiz, M. Z., Prince, S. M., Howard, T., Cogdell, R. J. & Isaacs, N. W. The Structure and Thermal Motion of the B800–B850 LH2 Complex from *Rps.acidophila* at 2.0Å Resolution and 100K: New Structural Features and Functionally Relevant Motions. *J. Mol. Biol.* **326,** 1523–1538 (2003).
83. Niwa, S. *et al.* Structure of the LH1–RC complex from Thermochromatium tepidum at 3.0 Å. *Nature* **508,** 228–232 (2014).
84. Jungas, C., Ranck, J. L., Rigaud, J. L., Joliot, P. & Verméglio, A. Supramolecular organization of the photosynthetic apparatus of *Rhodobacter sphaeroides*. *EMBO J.* **18,** 534–542 (1999).
85. Holden-Dye, K., Crouch, L. I. & Jones, M. R. Structure, function and interactions of the PufX protein. *Biochim. Biophys. Acta - Bioenerg.* **1777,** 613–630 (2008).
86. Jones, M. R. *et al.* Mutants of *Rhodobacter sphaeroides* lacking one or more pigment-protein complexes and complementation with reaction-centre, LH1, and LH2 genes. *Mol. Microbiol.* **6,** 1173–1184 (1992).
87. Jones, M. R., Visschers, R. W., van Grondelle, R. & Hunter, C. N. Construction and characterization of a mutant of *Rhodobacter sphaeroides* with the reaction center as the sole pigment-protein complex. *Biochemistry* **31,** 4458–4465 (1992).
88. Xu, P., Roy, L. M. & Croce, R. Functional organization of Photosystem II antenna complexes: CP29 under the spotlight. *Biochim. Biophys. Acta - Bioenerg.* **1858,** 815–822 (2017).
89. Dall'Osto, L., Ünlü, C., Cazzaniga, S. & van Amerongen, H. Disturbed excitation energy transfer in *Arabidopsis thaliana* mutants lacking minor antenna complexes of Photosystem II. *Biochim. Biophys. Acta - Bioenerg.* **1837,** 1981–1988 (2014).
90. Le Quiniou, C., van Oort, B., Drop, B., van Stokkum, I. H. M. & Croce, R. The high efficiency of Photosystem I in the green alga *Chlamydomonas reinhardtii* is maintained after the antenna size is substantially increased by the association of light-harvesting complexes II. *J. Biol. Chem.* **290,** 30587–30595 (2015).

91. Friebe, V. M. & Frese, R. N. Photosynthetic reaction center-based biophotovoltaics. *Curr. Opin. Electrochem.* **5,** 126–134 (2017).
92. Sener, M., Strumpfer, J., Singharoy, A., Hunter, C. N. & Kingdom, U. Overall energy conversion efficiency of a photosynthetic vesicle. *Elife* **5,** e09541 (2016).
93. Pan, X. *et al.* Structure of the maize photosystem I supercomplex with light-harvesting complexes I and II. *Science* **1113,** 1109–1113 (2018).
94. Wientjes, E., van Amerongen, H. & Croce, R. LHCII is an antenna of both photosystems after long-term acclimation. *Biochim. Biophys. Acta - Bioenerg.* **1827,** 420–426 (2013).
95. Tucker, J. D. *et al.* Membrane invagination in *Rhodobacter sphaeroides* is initiated at curved regions of the cytoplasmic membrane , then forms both budded and fully detached spherical vesicles. *Mol. Microbiol.* **76,** 833–847 (2010).
96. Adams, P. G. & Hunter, C. N. Adaptation of intracytoplasmic membranes to altered light intensity in *Rhodobacter sphaeroides*. *Biochim. Biophys. Acta - Bioenerg.* **1817,** 1616–1627 (2012).
97. Olsen, J. D., Hunter, C. N. & Schulten, K. Atomic-level structural and functional model of a bacterial photosynthetic membrane vesicle. *Proc Natl Acad Sci U S A* **104,** 15723–15728 (2007).
98. Lundell, D. J., Yamanaka, G. & Glazer, A. N. A Terminal Energy Acceptor of the Phycobilisome : The 75,000-dalton Polypeptide of *Synechococcus* 6301 Phycobilisomes-A New Biliprotein. *J. Cell Biol.* **91,** 315–319 (1981).
99. McConnell, M. D., Koop, R., Vasil, S. & Bruce, D. Regulation of the Distribution of Chlorophyll and Phycobilin-Absorbed Excitation Energy in Cyanobacteria . A Structure-Based Model for the Light State Transition 1. *Plant Physiol.* **130,** 1201–1212 (2002).
100. Saer, R. G. & Blankenship, R. E. Light harvesting in phototrophic bacteria: structure and function. *Biochem. J.* **474,** 2107–2131 (2017).
101. Olson, J. M. *et al.* The FMO protein. *Photosynth. Res.* **80,** 181–187 (2004).
102. Duan, S. *et al.* Bilayer Chlorophyll-Based Biosolar Cells Inspired from the Z-Scheme Process of Oxygenic Photosynthesis. *ACS Energy Lett.* **3,** 1708–1712 (2018).
103. Badura, A., Kothe, T., Schuhmann, W. & Rögner, M. Wiring photosynthetic enzymes to electrodes. *Energy Environ. Sci.* **4,** 3263–3274 (2011).
104. Boghossian, A. A., Ham, M.-H., Choi, J. H. & Strano, M. S. Biomimetic strategies for solar energy conversion: a technical perspective. *Energy Environ. Sci.* **4,** 3834–3843 (2011).
105. Wang, F., Liu, X. & Willner, I. Integration of photoswitchable proteins, photosynthetic reaction centers and semiconductor/biomolecule hybrids with electrode supports for optobioelectronic applications. *Adv. Mater.* **25,** 349–377 (2013).
106. Kim, Y., Shin, S. A., Lee, J., Yang, K. D. & Nam, K. T. Hybrid system of semiconductor and photosynthetic protein. *Nanotechnology* **25,** 1–20 (2014).
107. Nguyen, K. & Bruce, B. D. Growing green electricity: Progress and strategies for use of Photosystem i for sustainable photovoltaic energy conversion. *Biochim. Biophys. Acta - Bioenerg.* **1837,** 1553–1566 (2014).
108. Yehezkeli, O. *et al.* Integrated Photosystem II-based photo-bioelectrochemical cells. *Nat. Commun.* **3,** 742 (2012).
109. Ravi, S. K. & Tan, S. C. Progress and perspectives in exploiting photosynthetic biomolecules for solar energy harnessing. *Energy Environ. Sci.* **8,** 2551–2573 (2015).
110. Ravi, S. K., Udayagiri, V. S., Suresh, L. & Tan, S. C. Emerging Role of the Band-Structure Approach in Biohybrid Photovoltaics: A Path Beyond Bioelectrochemistry. *Adv. Funct. Mater.* **28,** 1–18 (2017).
111. Chatzipetrou, M. *et al.* Functionalization of gold screen printed electrodes with bacterial photosynthetic reaction centers by laser printing technology for mediatorless herbicide biosensing. *Electrochem. Commun.* **64,** 46–50 (2016).
112. Szabó, T. *et al.* Sensing photosynthetic herbicides in an electrochemical flow cell. *Photosynth. Res.* **132,** 127–134 (2017).

113. Swainsbury, D. J. K., Friebe, V. M., Frese, R. N. & Jones, M. R. Evaluation of a biohybrid photoelectrochemical cell employing the purple bacterial reaction centre as a biosensor for herbicides. *Biosens. Bioelectron.* **58,** 172–178 (2014).
114. Terasaki, N. *et al.* Bio-photosensor: Cyanobacterial Photosystem I coupled with transistor via molecular wire. *Biochim. Biophys. Acta - Bioenerg.* **1767,** 653–659 (2007).
115. Ravi, S. K. *et al.* Photosynthetic Bioelectronic Sensors for Touch Perception, UV-Detection, and Nanopower Generation: Toward Self-Powered E-Skins. *Adv. Mater.* **1802290,** 1–9 (2018).
116. Ravi, S. K. *et al.* Photosynthetic apparatus of *Rhodobacter sphaeroides* exhibits prolonged charge storage. *Nat. Commun.* **10,** 902 (2019).
117. Krassen, H. *et al.* Photosynthetic hydrogen production by a hybrid complex of photosystem I and [NiFe]-hydrogenase. *ACS Nano* **3,** 4055–4061 (2009).
118. Oliver, J. W. K. & Atsumi, S. Metabolic design for cyanobacterial chemical synthesis. *Photosynth. Res.* **120,** 249–261 (2014).
119. Efrati, A. *et al.* Assembly of photo-bioelectrochemical cells using Photosystem I-functionalized electrodes. *Nat. Energy* **1,** 15021 (2016).
120. Sokol, K. P. *et al.* Bias-free photoelectrochemical water splitting with Photosystem II on a dye-sensitized photoanode wired to hydrogenase. *Nat. Energy* **3,** 944–951 (2018).
121. Jordan, P. *et al.* Three-dimensional structure of cyanobacterial Photosystem I at 2.5 Å resolution. *Nature* **411,** 909–917 (2001).
122. Grotjohann, I. & Fromme, P. Structure of cyanobacterial Photosystem I. *Photosynth. Res.* **85,** 51–72 (2005).
123. Rappaport, F., Guergova-Kuras, M., Nixon, P. J. & Diner, B. A. Kinetics and Pathways of Charge Recombination in Photosystem II. *Biochemistry* **41,** 8518–8527 (2002).
124. Kornienko, N. *et al.* Oxygenic Photoreactivity in Photosystem II Studied by Rotating Ring Disk Electrochemistry. *J. Am. Chem. Soc.* **140,** 17923–17931 (2018).
125. Blankenship, R. E. *et al.* Comparing photosynthetic and photovoltaic efficiencies and recognizing the potential for improvement. *Science* **332,** 805–809 (2011).
126. Blankenship, R. E. *Molecular Mechanisms of Photosynthesis*. *Molecular Mechanisms of Photosynthesis* (2008). doi:10.1002/9780470758472
127. Ort, D. R. *et al.* Redesigning photosynthesis to sustainably meet global food and bioenergy demand. *Proc. Natl. Acad. Sci.* **112,** 8529–8536 (2015).
128. Kromdijk, J. *et al.* Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science (80-.).* **354,** 857–861 (2016).
129. South, P. F., Cavanagh, A. P., Liu, H. W. & Ort, D. R. Synthetic glycolate metabolism pathways stimulate crop growth and productivity in the field. *Science* **363,** eaat9077 (2019).
130. Shen, B. *et al.* Engineering a New Chloroplastic Photorespiratory Bypass to Increase Photosynthetic Efficiency and Productivity in Rice. *Mol. Plant* **12,** 199–214 (2019).
131. Gundlach, K., Werwie, M., Wiegand, S. & Paulsen, H. Filling the 'green gap' of the major light-harvesting chlorophyll a/b complex by covalent attachment of Rhodamine Red. *Biochim. Biophys. Acta - Bioenerg.* **1787,** 1499–1504 (2009).
132. Springer, J. W. *et al.* Biohybrid photosynthetic antenna complexes for enhanced light-harvesting. *J. Am. Chem. Soc.* **134,** 4589–4599 (2012).
133. Grayson, K. J. *et al.* Augmenting light coverage for photosynthesis through YFP-enhanced charge separation at the *Rhodobacter sphaeroides* reaction centre. *Nat. Commun.* **8,** 1–12 (2017).
134. Ravi, S. K. *et al.* A Mechanoresponsive Phase-Changing Electrolyte Enables Fabrication of High-Output Solid-State Photobioelectrochemical Devices from Pigment-Protein Multilayers. *Adv. Mater.* **30,** 1–8 (2018).
135. Liu, J. *et al.* Engineered photoproteins that give rise to photosynthetically-incompetent bacteria are effective as photovoltaic materials for biohybrid photoelectrochemical cells. *Faraday Discuss.* **207,** 307–327 (2018).

136. Ravi, S. K. *et al.* Enhanced Output from Biohybrid Photoelectrochemical Transparent Tandem Cells Integrating Photosynthetic Proteins Genetically Modified for Expanded Solar Energy Harvesting. *Adv. Energy Mater.* **7,** 1–7 (2017).
137. Tan, S. C. *et al.* Superhydrophobic carbon nanotube electrode produces a near-symmetrical alternating current from photosynthetic protein-based photoelectrochemical cells. *Adv. Funct. Mater.* **23,** 5556–5563 (2013).
138. Tan, S. C., Crouch, L. I., Mahajan, S., Jones, M. R. & Welland, M. E. Increasing the open-circuit voltage of photoprotein-based photoelectrochemical cells by manipulation of the vacuum potential of the electrolytes. *ACS Nano* **6,** 9103–9109 (2012).
139. Tan, S. C., Crouch, L. I., Jones, M. R. & Welland, M. Generation of alternating current in response to discontinuous illumination by photoelectrochemical cells based on photosynthetic proteins. *Angew. Chemie - Int. Ed.* **51,** 6667–6671 (2012).
140. Friebe, V. M. *et al.* Plasmon-Enhanced Photocurrent of Photosynthetic Pigment Proteins on Nanoporous Silver. *Adv. Funct. Mater.* **26,** 285–292 (2016).
141. Kamran, M. *et al.* Demonstration of asymmetric electron conduction in pseudosymmetrical photosynthetic reaction centre proteins in an electrical circuit. *Nat. Commun.* **6,** 6530 (2015).
142. Den Hollander, M. J. *et al.* Enhanced photocurrent generation by photosynthetic bacterial reaction centers through molecular relays, light-harvesting complexes, and direct protein-gold interactions. *Langmuir* **27,** 10282–10294 (2011).
143. Białek, R., Swainsbury, D. J. K., Wiesner, M., Jones, M. R. & Gibasiewicz, K. Modelling of the cathodic and anodic photocurrents from *Rhodobacter sphaeroides* reaction centres immobilized on titanium dioxide. *Photosynth. Res.* **138,** 103–114 (2018).
144. Paulsen, H., Rümler, U. & Rüdiger, W. Reconstitution of pigment-containing complexes from light-harvesting chlorophyll *a*/*b*-binding protein overexpressed in *Escherichia coli*. *Planta* **181,** 204–211 (1990).
145. Croce, R., Weiss, S. & Bassi, R. Carotenoid-binding Sites of the Major Light-harvesting Complex II of Higher Plants. *J. Biol. Chem.* **274,** 29613–29623 (1999).
146. Thomas, F., Boyle, A. L., Burton, A. J. & Woolfson, D. N. A set of *de novo* designed parallel heterodimeric coiled coils with quantified dissociation constants in the micromolar to sub-nanomolar regime. *J. Am. Chem. Soc.* **135,** 5161–5166 (2013).
147. Zakeri, B. *et al.* Peptide tag forming a rapid covalent bond to a protein, through engineering a bacterial adhesin. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 690–697 (2012).
148. Bromley, E. H. C., Sessions, R. B., Thomson, A. R. & Woolfson, D. N. Designed alpha-helical tectons for constructing multicomponent synthetic biological systems. *J. Am. Chem. Soc.* **131,** 928–930 (2009).
149. Fierer, J. O., Veggiani, G. & Howarth, M. SpyLigase peptide-peptide ligation polymerizes affibodies to enhance magnetic cancer cell capture. *Proc. Natl. Acad. Sci. U. S. A.* **111,** E1176-81 (2014).
150. Ji, X., Wang, W. & Mattoussi, H. Controlling the spectroscopic properties of quantum dots via energy transfer and charge transfer interactions: Concepts and applications. *Nano Today* **11,** 98–121 (2015).
151. Werwie, M., Fehr, N., Xu, X., Basché, T. & Paulsen, H. Comparison of quantum dot-binding protein tags: Affinity determination by ultracentrifugation and FRET. *Biochim. Biophys. Acta - Gen. Subj.* **1840,** 1651–1656 (2014).
152. Medintz, I. L., Uyeda, H. T., Goldman, E. R. & Mattoussi, H. Quantum dot bioconjugates for imaging, labelling and sensing. *Nat. Mater.* **4,** 435–446 (2005).
153. Nozik, A. J. Quantum dot solar cells. *Phys. E Low-Dimensional Syst. Nanostructures* **14,** 115–120 (2002).
154. Medintz, I. L. & Mattoussi, H. Quantum dot-based resonance energy transfer and its growing application in biology. *Phys. Chem. Chem. Phys.* **11,** 17–45 (2009).
155. Efros, A. L. & Rosen, M. The Electronic Structure of Semiconductor Nanocrystals. *Annu. Rev. Mater. Sci.* **30,** 475–521 (2000).
156. Li, L., Fierer, J. O., Rapoport, T. A. & Howarth, M. Structural analysis and optimization

of the covalent association between SpyCatcher and a peptide tag. *J. Mol. Biol.* **426,** 309–317 (2014).
157. Veggiani, G. *et al.* Programmable polyproteams built using twin peptide superglues. *Proc. Natl. Acad. Sci.* **113,** 1202–1207 (2016).
158. Lupas, A. N. & Gruber, M. The structure of alpha-helical coiled coils. *Adv. Protein Chem.* **70,** 37–78 (2005).
159. Woolfson, D. N., Bartlett, G. J., Bruning, M. & Thomson, A. R. New currency for old rope: From coiled-coil assemblies to alpha-helical barrels. *Curr. Opin. Struct. Biol.* **22,** 432–441 (2012).
160. Wood, C. W. *et al.* ISAMBARD: An open-source computational environment for biomolecular analysis, modelling and design. *Bioinformatics* **33,** 3043–3050 (2017).
161. Wu, Q. *et al.* Conserved residues that modulate protein trans -splicing of Npu DnaE split intein. *Biochem. J* **255,** 247–255 (2014).
162. Myszka, D. G. & Chaiken, I. M. Design and characterization of an intramolecular antiparallel coiled coil peptide. *Biochemistry* **33,** 2363–2372 (1994).
163. Thomas, F., Niitsu, A., Oregioni, A., Bartlett, G. J. & Woolfson, D. N. Conformational Dynamics of Asparagine at Coiled-Coil Interfaces. *Biochemistry* **56,** 6544–6554 (2017).
164. Lee, M. J. *et al.* Engineered synthetic scaffolds for organizing proteins within the bacterial cytoplasm. *Nat. Chem. Biol.* **14,** 142–147 (2018).
165. Lee, M. J. *et al. De novo* targeting to the cytoplasmic and luminal side of bacterial microcompartments. *Nat. Commun.* **9,** 3413 (2018).
166. Stevens, A. J. *et al.* Design of a Split Intein with Exceptional Protein Splicing Activity. *J. Am. Chem. Soc.* **138,** 2162–2165 (2016).
167. Chong, S. *et al.* Protein splicing involving the Saccharomyces cerevisiae VMA intein: The steps in the splicing pathway, side reactions leading to protein cleavage, and establishment of an in vitro splicing system. *J. Biol. Chem.* **271,** 22159–22168 (1996).
168. Bethesda Research Laboratories. BRL pUC host: E. coli DH5αTM competent cells. *Bethesda Res. Lab. Focus* **8,** 9–12 (1986).
169. Simon, R., Priefer, U. & Pühler, A. A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. *Bio/Technology* **1,** 784–791 (1983).
170. Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189,** 113–130 (1986).
171. Robichon, C., Luo, J., Causey, T. B., Benner, J. S. & Samuelson, J. C. Engineering *Escherichia coli* BL21(DE3) Derivative Strains To Minimize *E. coli* Protein Contamination after Purification by Immobilized Metal Affinity Chromatography. *Appl. Environ. Microbiol.* **77,** 4634–4646 (2011).
172. Miroux, B. & Walker, J. E. Over-production of Proteins in *Escherichia coli*: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. *J. Mol. Biol.* **260,** 289–298 (1996).
173. Wagner, S. *et al.* Tuning *Escherichia coli* for membrane protein overexpression. *Proc. Natl. Acad. Sci.* **105,** 14371–14376 (2008).
174. Novy, R., Drott, D., Yaeger, K. & Mierendorf, R. Overcoming the codon bias of E. coli for enhanced protein expression. *Innovations* **12,** 1–3 (2001).
175. Hunter, C. N. & Turner, G. Transfer of Genes Coding for Apoproteins of Reaction Centre and Light-harvesting LH1 Complexes to *Rhodobacter sphaeroides*. *Microbiology* **134,** 1471–1480 (1988).
176. Sistrom, W. R. Transfer of chromosomal genes mediated by plasmid R68.45 in *Rhodopseudomonas sphaeroides*. *J. Bacteriol.* **131,** 526–532 (1977).
177. McAuley-Hecht, K. E. *et al.* Structural Studies of Wild-Type and Mutant Reaction Centers from an Antenna-Deficient Strain of Rhodobacter sphaeroides: Monitoring the Optical Properties of the Complex from Bacterial Cell to Crystal. *Biochemistry* **37,** 4740–4750 (1998).
178. Alderton, D. *et al.* A versatile ligation-independent cloning method suitable for high-

throughput expression screening applications. *Nucleic Acids Res.* **35,** e45–e45 (2007).
179. Hobe, S., Prytulla, S., Kuhlbrandt, W. & Paulsen, H. Trimerization and crystallization of reconstituted light-harvesting chlorophyll *a/b* complex. *EMBO J.* **13,** 3423–3429 (1994).
180. Schmid, V. H., Cammarata, K. V, Bruns, B. U. & Schmidt, G. W. *In vitro* reconstitution of the photosystem I light-harvesting complex LHCI-730: heterodimerization is required for antenna pigment organization. *Proc. Natl. Acad. Sci. U. S. A.* **94,** 7667–7672 (1997).
181. Porra, R. J., Thompson, W. A. & Kriedemann, P. E. Determination of Accurate Extinction Coefficients and Simultaneous-Equations for Assaying Chlorophyll-*a* and Chlorophyll-*b* Extracted with 4 Different Solvents - Verification of the Concentration of Chlorophyll Standards by Atomic-Absorption S. *Biochim. Biophys. Acta* **975,** 384–394 (1989).
182. Wei, X. *et al.* Structure of spinach Photosystem II-LHCII supercomplex at 3.2 Å resolution. *Nature* **534,** 69–74 (2016).
183. Ma, J. & Xia, D. The use of blue native PAGE in the evaluation of membrane protein aggregation states for crystallization. *J. Appl. Crystallogr.* **41,** 1150–1160 (2008).
184. Würth, C., Grabolle, M., Pauli, J., Spieles, M. & Resch-Genger, U. Relative and absolute determination of fluorescence quantum yields of transparent samples. *Nat. Protoc.* **8,** 1535–1350 (2013).
185. Gardecki, J. A. & Maroncelli, M. Set of secondary emission standards for calibration of the spectral responsivity in emission spectroscopy. *Appl. Spectrosc.* **52,** 1179–1189 (1998).
186. Rurack, K. & Spieles, M. Fluorescence quantum yields of a series of red and near-infrared dyes emitting at 600-1000 nm. *Anal. Chem.* **83,** 1232–1242 (2011).
187. Kruger, T. P. J. *et al.* Disentangling the low-energy states of the major light-harvesting complex of plants and their role in photoprotection. *Biochim. Biophys. Acta - Bioenerg.* **1837,** 1027–1038 (2014).
188. Krivov, G. G., Shapovalov, M. V & Dunbrack Jr., R. L. Improved prediction of protein side-chain conformations with SCWRL4. *Proteins Struct. Funct. Bioinforma.* **77,** 778–795 (2009).
189. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* **15,** 5.6.1-5.6.30 (2016).
190. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2,** 19–25 (2015).
191. Lipfert, J., Columbus, L., Chu, V. B., Lesley, S. A. & Doniach, S. Size and shape of detergent micelles determined by small-angle X-ray scattering. *J. Phys. Chem. B* **111,** 12427–12438 (2007).
192. Briand, G. & Dauchot, O. Crystallization of Self-Propelled Hard Discs. *Phys. Rev. Lett.* **117,** 098004 (2016).
193. Mirvakili, S. M. *et al.* Photoactive electrodes incorporating electrosprayed bacterial reaction centers. *Adv. Funct. Mater.* **24,** 4789–4794 (2014).
194. Yaghoubi, H. *et al.* Large Photocurrent Response and External Quantum Efficiency in Biophotoelectrochemical Cells Incorporating Reaction Center Plus Light Harvesting Complexes. *Biomacromolecules* **16,** 1112–1118 (2015).
195. Vos, M. H., Breton, J. & Martin, J.-L. Electronic energy transfer within the hexamer cofactor system of bacterial reaction centers. *J. Phys. Chem. B* **101,** 9820–9832 (1997).
196. Jordanides, X. J., Scholes, G. D. & Fleming, G. R. The Mechanism of Energy Transfer in the Bacterial Photosynthetic Reaction Center. *J. Phys. Chem. B* **105,** 1652–1669 (2001).
197. Clegg, R. M. Fluorescence resonance energy transfer and nucleic acids. *Methods Enzymol.* **211,** 353–388 (1992).
198. Pons, T., Medintz, I. I. L., Wang, X., English, D. S. & Mattoussi, H. Solution-phase

single quantum dot fluorescence resonance energy transfer. *J. Am. Chem. Soc.* **128,** 15324–15331 (2006).
199. Govorov, A. O. Enhanced optical properties of a photosynthetic system conjugated with semiconductor nanoparticles: The role of Förster transfer. *Adv. Mater.* **20,** 4330–4335 (2008).
200. Martinez, J. C. *et al.* in *Applications of Calorimetry in a Wide Context* (eds. Cobos, E. S. & Elkordy, A. A.) 73–104 (InTech, 2013). doi:10.5772/53311
201. Ridge, J. P., Van Brederode, M. E., Goodwin, M. G., Van Grondelle, R. & Jones, M. R. Mutations that modify or exclude binding of the Q_A ubiquinone and carotenoid in the reaction center from *Rhodobacter sphaeroides*. *Photosynth. Res.* **59,** 9–26 (1999).
202. Jackson, J. A. *et al.* Energy Transfer in *Rhodobacter sphaeroides* Reaction Centers with the Initial Electron Donor Oxidized or Missing. *J. Phys. Chem. B* **101,** 5747–5754 (1997).
203. Kirmaier, C., Gaul, D., DeBey, R., Holten, D. & Schenck, C. Charge separation in a reaction center incorporating bacteriochlorophyll for photoactive bacteriopheophytin. *Science* **251,** 922–927 (1991).
204. Watson, A. J. *et al.* Replacement or exclusion of the B-branch bacteriopheophytin in the purple bacterial reaction centre: The H_B cofactor is not required for assembly or core function of the *Rhodobacter sphaeroides* complex. *Biochim. Biophys. Acta - Bioenerg.* **1710,** 34–46 (2005).
205. Miller, W. G. & Goebel, C. V. Dimensions of Protein Random Coils. *Biochemistry* **7,** 3925–3935 (1968).
206. Operamolla, A. *et al.* “Garnishing” the photosynthetic bacterial reaction center for bioelectronics. *J. Mater. Chem. C* **3,** 6471–6478 (2015).
207. Milano, F. *et al.* Enhancing the light harvesting capability of a photosynthetic reaction center by a tailored molecular fluorophore. *Angew. Chemie - Int. Ed.* **51,** 11019–11023 (2012).
208. Hassan Omar, O. *et al.* Synthetic Antenna Functioning As Light Harvester in the Whole Visible Region for Enhanced Hybrid Photosynthetic Reaction Centers. *Bioconjug. Chem.* **27,** 1614–1623 (2016).
209. Dutta, P. K. *et al.* Reengineering the optical absorption cross-section of photosynthetic reaction centers. *J. Am. Chem. Soc.* **136,** 4599–4604 (2014).
210. Dutta, P. K. *et al.* A DNA-directed light-harvesting/reaction center system. *J. Am. Chem. Soc.* **136,** 16618–16625 (2014).
211. Nabiev, I. *et al.* Fluorescent quantum dots as artificial antennas for enhanced light harvesting and energy transfer to photosynthetic reaction centers. *Angew. Chemie - Int. Ed.* **49,** 7217–7221 (2010).
212. Lukashev, E. P., Knox, P. P., Oleinikov, I. P., Seifullina, N. K. & Grishanova, N. P. Investigation of Stability of Photosynthetic Reaction Center and Quantum Dot Hybrid Films. *Biochemistry* **81,** 58–63 (2016).
213. Maksimov, E. G., Lukashev, E. P., Seifullina, N. K., Nizova, G. V. & Pashchenko, V. Z. Photophysical Properties of Hybrid Complexes of Quantum Dots and Reaction Centers of Purple Photosynthetic Bacteria *Rhodobacter Sphaeroides* Adsorbed on Crystalline Mesoporous TiO2 Films. *Nanotechnol. Russ.* **8,** 423–431 (2013).
214. Campillo, A. J., Hyer, R. C., Monger, T. G., Parson, W. W. & Shapiro, S. L. Light collection and harvesting processes in bacterial photosynthesis investigated on a picosecond time scale. *Proc. Natl. Acad. Sci. U. S. A.* **74,** 1997–2001 (1977).
215. Valkunas, L., Trinkunas, G., Liuolia, V. & van Grondelle, R. Nonlinear annihilation of excitations in photosynthetic systems. *Biophys. J.* **69,** 1117–1129 (1995).
216. Pflock, T. J. *et al.* The electronically excited states of LH2 complexes from *Rhodopseudomonas acidophila* Strain 10050 studied by time-resolved spectroscopy and dynamic Monte Carlo simulations. II. Homo-arrays of LH2 complexes reconstituted into phospholipid model membr. *J. Phys. Chem. B* **115,** 8821–8831 (2011).
217. Sumino, A. *et al.* Influence of phospholipid composition on self-assembly and energy-

transfer efficiency in networks of light-harvesting 2 complexes. *J. Phys. Chem. B* **117,** 10395–10404 (2013).
218. Sapsford, K. E. *et al.* Kinetics of metal-affinity driven self-assembly between proteins or peptides and CdSe-ZnS quantum dots. *J. Phys. Chem. C* **111,** 11528–11538 (2007).
219. Timpmann, K. *et al.* Efficiency of light harvesting in a photosynthetic bacterium adapted to different levels of light. *BBA - Bioenerg.* **1837,** 1835–1846 (2014).
220. Scholes, G. D., Fleming, G. R., Olaya-Castro, A. & van Grondelle, R. Lessons from nature about solar light harvesting. *Nat. Chem.* **3,** 763–774 (2011).
221. Croce, R. & van Amerongen, H. Natural strategies for photosynthetic light harvesting. *Nat. Chem. Biol.* **10,** 492–501 (2014).
222. Milano, F., Punzi, A., Ragni, R., Trotta, M. & Farinola, G. M. Photonics and Optoelectronics with Bacteria : Making Materials from Photosynthetic Microorganisms. *Adv. Funct. Mater.* **1805521,** 1–17 (2018).
223. Kornienko, N., Zhang, J. Z., Sakimoto, K. K., Yang, P. & Reisner, E. Interfacing nature's catalytic machinery with synthetic materials for semi-artificial photosynthesis. *Nat. Nanotechnol.* **13,** 890–899 (2018).
224. Liu, J., Mantell, J., Di Bartolo, N. & Jones, M. R. Mechanisms of Self-Assembly and Energy Harvesting in Tuneable Conjugates of Quantum Dots and Engineered Photovoltaic Proteins. *Small* **1804267,** 1–16 (2018).
225. Novoderezhkin, V., Marin, A. & van Grondelle, R. Intra- and inter-monomeric transfers in the light harvesting LHCII complex: The Redfield-Förster picture. *Phys. Chem. Chem. Phys.* **13,** 17093–17103 (2011).
226. Pagano, A., Cinque, G. & Bassi, R. In vitro reconstitution of the recombinant photosystem II light-harvesting complex CP24 and its spectroscopic characterization. *J. Biol. Chem.* **273,** 17154–17165 (1998).
227. Caffarri, S., Croce, R., Cattivelli, L. & Bassi, R. A look within LHCII: Differential analysis of the Lhcb1-3 complexes building the major trimeric antenna complex of higher-plant photosynthesis. *Biochemistry* **43,** 9467–9476 (2004).
228. Sun, L. *et al.* Crystal structure of a bacterial homologue of glucose transporters GLUT1–4. *Nature* **490,** 361–366 (2012).
229. Kouřil, R., Wientjes, E., Bultema, J. B., Croce, R. & Boekema, E. J. High-light vs. low-light: Effect of light acclimation on photosystem II composition and organization in *Arabidopsis thaliana*. *Biochim. Biophys. Acta - Bioenerg.* **1827,** 411–419 (2013).
230. Werwie, M. *et al.* Light-harvesting chlorophyll protein (LHCII) drives electron transfer in semiconductor nanocrystals. *BBA - Bioenerg.* **1859,** 174–181 (2018).
231. Wientjes, E., van Amerongen, H. & Croce, R. Quantum yield of charge separation in Photosystem II: Functional effect of changes in the antenna size upon light acclimation. *J. Phys. Chem. B* **117,** 11200–11208 (2013).
232. Owens, T. G., Webb, S. P., Mets, L., Alberte, R. S. & Fleming, G. R. Antenna size dependence of fluorescence decay in the core antenna of Photosystem I : Estimates of charge separation and energy transfer rates. *Proc Natl Acad Sci U S A* **84,** 1532–1536 (1987).
233. Hohmann-Marriott, M. F. & Blankenship, R. E. Evolution of photosynthesis. *Annu Rev Plant Biol* **62,** 515–548 (2011).
234. Björkman, O. & Demmig, B. Photon yield of O2 evolution and chlorophyll fluorescence characteristics at 77 K among vascular plants of diverse origins. *Planta.* **170,** 489–504 (1987).
235. Scheuring, S. & Sturgis, J. N. Chromatic adaptation of photosynthetic membranes. *Science (80-.).* **309,** 484–487 (2005).
236. Ilioaia, C., Johnson, M. P., Horton, P. & Ruban, A. V. Induction of efficient energy dissipation in the isolated light-harvesting complex of Photosystem II in the absence of protein aggregation. *J. Biol. Chem.* **283,** 29505–29512 (2008).
237. Kolber, Z. S. *et al.* Contribution of aerobic photoheterotrophic bacteria to the carbon cycle in the ocean. *Science* **292,** 2492–2495 (2001).

238. Yehezkeli, O., Tel-Vered, R., Michaeli, D., Willner, I. & Nechushtai, R. Photosynthetic reaction center-functionalized electrodes for photo-bioelectrochemical cells. *Photosynth. Res.* **120,** 71–85 (2014).
239. Van Bezouwen, L. S. *et al.* Subunit and chlorophyll organization of the plant Photosystem II supercomplex. *Nat. Plants* **3,** 1–11 (2017).
240. Passarini, F., Wientjes, E., van Amerongen, H. & Croce, R. Photosystem I light-harvesting complex Lhca4 adopts multiple conformations: Red forms and excited-state quenching are mutually exclusive. *Biochim. Biophys. Acta - Bioenerg.* **1797,** 501–508 (2010).
241. Gibasiewicz, K. *et al.* Mechanism of recombination of the $P^+H_A^-$radical pair in mutant *Rhodobacter sphaeroides* reaction centers with modified free energy gaps between $P^+B_A^-$ and $P^+H_A^-$. *J. Phys. Chem. B* **115,** 13037–13050 (2011).
242. Buldun, C. M., Jean, J. X., Bedford, M. R. & Howarth, M. SnoopLigase Catalyzes Peptide-Peptide Locking and Enables Solid-Phase Conjugate Isolation. *J. Am. Chem. Soc.* **140,** 3008–3018 (2018).
243. Swainsbury, D. J. K. *et al.* Directed assembly of defined oligomeric photosynthetic reaction centres through adaptation with programmable extra-membrane coiled-coil interfaces. *Biochim. Biophys. Acta - Bioenerg.* **1857,** 1829–1839 (2016).
244. Fletcher, J. M. *et al.* A Basis Set of de novo Coiled-Coil Peptide Oligomers for Rational Protein Design and Synthetic Biology. *ACS Synth. Biol.* **1,** 240–250 (2012).
245. Standfuss, J. & Kühlbrandt, W. The three isoforms of the light-harvesting complex II: Spectroscopic features, trimer formation, and functional roles. *J. Biol. Chem.* **279,** 36884–36891 (2004).
246. Opacic, M., Durand, G., Bosco, M., Polidori, A. & Popot, J. L. Amphipols and Photosynthetic Light-Harvesting Pigment-Protein Complexes. *J. Membr. Biol.* **247,** 1031–1041 (2014).
247. Marsden, H. R., Korobko, A. V., Zheng, T., Voskuhl, J. & Kros, A. Controlled liposome fusion mediated by SNARE protein mimics. *Biomater. Sci.* **1,** 1046 (2013).

Appendices

Appendix 2.1(A). 10 x M22+ concentrate

	per 4 L
KH_2PO_4	122.4 g
K_2HPO_4	92 g
Sodium lactate solution (lactic acid)	100 mL
$(NH_4)_2SO_4$	20 g
NaCl	20 g
Solution C	800 mL
Sodium succinate	173.7 g
Sodium glutamate (glutamic acid)	10.8 g
Aspartic acid	1.6 g

Bring to pH 6.8 with KOH or NaOH). Autoclave at 15 psi for 20 mins.

Appendix 2.1(B). Solution C

	per 4 L
Nitrilotriacetic acid	40 g
$MgCl_2$	96 g
$CaCl_2$	13.36 g
EDTA	500 mg
$ZnCl_2$	1044 mg
$FeCl_2$	1000 mg
$MnCl_2$	360 mg
$(NH_4)6Mo_7O_{27}.4H_2O$	37 mg
$CuCl_2$	31 mg
$Co(NO_3)_2$	49.6 mg
(Ortho)boric acid	22.8 mg

Bring pH to between 6.8 and 6.9 with KOH (or NaOH).
Store in 800 mL aliquots at –20°C.

Appendix 2.1(C). 1 x M22+ liquid and solid medium

Per litre, dilute 100 mL 10 x M22+ concentrate in 900 mL deionised water. Add 1.0 g tryptone. For solid medium add 15 $g.L^{-1}$ agar. Autoclave at 15 psi for 20 mins.

Appendix 2.2. Vitamins solution

	per 100 mL
Nicotinic acid	1 g
Thiamine	0.5 g
Paraminobenzoic acid	0.1 g
Biotin	0.01 g

Filter sterilize and freeze at -20°C. Use at 200 $\mu L.L^{-1}$

Appendix 2.3. Luria Bertani liquid medium

	per 1 L
Tryptone	10 g
NaCl	10 g
Yeast extract	5 g

Adjust to pH 7.5 with NaOH. Autoclave at 15 psi for 20 mins. Add vitamins and antibiotics after cooling.

Appendix 2.4. Buffers for agarose gel electrophoresis

50x TAE stock

	per 1 L
Tris	242 g
Glacial acetic acid (100 %)	57.1 mL
EDTA	18.6 g

Bring pH to 8 with HCl

10x TBE stock

	per 1 L
Tris	270 g
Boric acid	137.5 g
EDTA	18.625 g

Bring pH to 8 with HCl

Appendix 2.5. RFI and RFII media

RFI	per 132 mL	
RbCl	100 mM	13.2 mL of 1 M stock
$MnCl_2$	50 mM	6.6 mL of 1 M stock
K acetate	30 mM	3.96 mL of 1 M stock
$CaCl_2$	10 mM	1.32 mL of 1 M stock
Glycerol	15% (w/v)	19.8 g

Bring pH to 5.4-5.6 with 0.2 M acetic acid. Filter sterilise. Store at 4°C.

RFII	per 32 mL	
MOPS	10 mM	0.64 mL of 1 M stock
RbCl	10 mM	0.32 mL of 1 M stock
$CaCl_2$	75 mM	2.4 mL of 1 M stock
Glycerol	15 % (w/v)	4.8 g

Bring pH to 6.8 with NaOH. Filter sterilise. Store at 4°C.